1979

Hierarchical Learning As A Function Of Concise Informational Feedback With Regard To Ability, Age, And Sex Of Identical Twins.

Ingrid A. Rimland
University of the Pacific

HIERARCHICAL LEARNING AS A FUNCTION OF CONCISE INFORMATIONAL FEEDBACK WITH REGARD TO ABILITY, AGE, AND SEX OF IDENTICAL TWINS

A Dissertation

Presented to

the Faculty of the Graduate School

University of the Pacific

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Education

by

Ingrid Rimland

June 1979

This dissertation, written and submitted by

Ingrid Rimland

is approved for recommendation to the Committee on Graduate Studies, University of the Pacific

Dean of the School or Department Chairman:

Oscar T. Jarvis

Oscar T. Jarvis, Dean

Dissertation Committee:

B. R. Hopkins Chairman

Dr. B.R. Hopkins

Roger L Reimer

Dr. Roger Reimer

Augustine García

Dr. Augustine García

Larry Lawson

Dr. Larry L. Lawson

D W Matheson

Dr. Douglas Matheson

Dated May 17, 1979

ABSTRACT

Originated by current demands for accountability in schools and resultant appeals for a return to the use of scholastic feedback, this study investigated the effects of feedback in light of certain characteristics of learning, and in light of student characteristics such as ability, age, and sex. The purpose of the study was to investigate the impact of specific informational feedback on hierarchical learning in a controlled laboratory setting, using the experimental controls inherent in identical twins. The Wechsler Intelligence Scale for Children/Revised (WISC-R) was used "unconventionally" as a series of discrete hierarchical learning tasks subject to concise experimental manipulation. Possible differences with regard to different forms of learning, as embodied in the scales of the instrument, were investigated. Possible differences with regard to ability, age, and sex were also examined, on the assumption that the results would shed light on potentially controversial practices tied to accountability, such as objective grading, ability grouping for instructional purposes, the use of merit systems, and others.

The study used a cross-sectional matched-subjects design, assigning 30 sets of identical twins randomly into an experimental group and a control group. The treatment was immediate feedback given in response to solved items, as specified in the WISC-R Manual. The underlying theory was that scales embodying a principle, rule, or "Gestalt" would lend themselves to feedback benefits, and that the benefits of knowing an "ideal" answer to lower order items would generalize into increased performance on higher order tasks. Student characteristics such as ability, age, and sex were also expected to be influenced differentially by feedback, but the direction was not specified.

The data were analyzed by means of a t-test and an analysis of variance. The findings showed the results of the effects of feedback to be task-specific as well as student characteristics-specific. The performance of the experimental group was significantly influenced in six out of ten WISC-R subscales and in all three summation scales, for the most part in a positive direction. The four scales that showed no difference when analyzed without regard to ability, age, and sex showed significant differences when further partitioned according to these student characteristics. It appeared from this study that bright children, on the average, gained more by informational feedback than slow children, and that old children gained more from feedback than young children in some tasks but not all. Sex was not found to be an important variable with regard to feedback except in two instances. The data further showed that those tasks most easily identified as "school related" (such as recall of facts or computational mastery) were negatively affected by feedback, while

less threatening tasks such as assembling block designs or giving conventional answers pertaining to social rules were positively affected. The overall findings were consistent with previous research showing that informational feedback is a potent and powerful modifier of learning, with beneficial results outbalancing negative results.

ACKNOWLEDGMENTS

Many friends and colleagues have helped in tangible as well as intangible ways. A special token of gratitude goes to Dr. Walter Quiring of Germany who was my friend, teacher, and spiritual mentor throughout my formative years. My friend Marta Pippin deserves mentioning--she cared and listened and understood. Of my committee members, a special "thank you" goes to my Chairman, Dr. Bobby Hopkins. Dr. Roger Reimer, Dr. Douglas Matheson, Dr. Larry Lawson, and Dr. Gus Garcia have graciously and generously shared their knowledge and their time. Dr. Lewis Aiken "rescued" me when computer phobia threatened to destroy me. My teenage son Rudy punched my cards, kept my car running, and cooked his own meals without a single word of protest. And last but not least, my outstanding and indefatigable typist, Mrs. Ruth Perry, kept every single deadline. My gratitude to all!

TABLE OF CONTENTS

LIST OF TABLES

Chapter 1

THE PROBLEM

INTRODUCTION

Current Societal Context

Evidence of continuous scholastic decline and resulting demands for accountability seem to abound.[1] The U.S. Office of Education, through the National Assessment of Educational Progress, reports that in the Spring of 1977, 13 percent of the nation's 17-year-olds were functionally illiterate. Elementary and secondary school students' scores have dropped for more than a decade. Verbal scores on the Scholastic Aptitude Test have seen a decline of 49 points between 1963 and 1977.[2] At the University of California at Berkeley, where students come from the top eight percent of California's high school graduates, nearly half the freshmen have been so deficient in recent years in their writing

---

[1]Robert L. Ebel, "The Case for Minimum Competency," Phi Delta Kappan, 59 (April, 1978), pp. 546-48.

[2]Focus 4: Learning to Read, Educational Testing Service, 1978.

ability as to warrant a remedial course dubbed "bonehead English."[3] At the City College of New York, where open admissions policies began in 1970, a staggering 90 percent of all students now take some form of remedial instruction.[4]

While academic scores have consistently declined for more than a decade, the cost of schooling has quadrupled.[5] The "California Tax Payers' Revolt," embedded in the easy passage of Proposition 13, was aimed in part at public schools, since it was widely felt that education was no longer as productive as might be expected.[6] Popham, a well-known spokesman for the accountability movement, has repeatedly called for precise evidence of actual student achievement, arguing that only by tying output to the characteristics of instructional input will schooling become more efficient.[7] In California, Stull has translated this demand into

---

[3]Focus 5: The Concern for Writing, Educational Testing Service, 1978.

[4]Ibid., pp. 2-3.

[5]CBS News: Report Card on American Education, narrated by Walter Cronkite and Charles Hollingsworth, August 22-24, 1978.

[6]Joan C. Baratz and Jay H. Moskowitz, "Proposition 13: How and Why It Happened," Phi Delta Kappan, 60 (September, 1978), pp. 9-11.

[7]W. James Popham, Educational Evaluation (Englewood Cliffs: Prentice Hall, Inc., 1975).

administratively exacting legislation.[8] The case of Peter Doe versus San Francisco Unified School District[9] has exemplified the fact that, in the absence of scholastic evidence, administrators might be faced with unpleasant legal consequences as well. More and more, accountability focuses on school site and, specifically, on demands for an efficient, output-oriented school administration.

In their search for improvement of schooling, some educational leaders are questioning causal connections that might impair efficiency.[10] Many see current educational failure as rooted in certain assumptions concerning the characteristics of learning, the nature of the learner, and the objectives that are to be mastered.[11] The characteristics of learning might entail a quest for psychological strategies that would increase efficiency. Learner variables might be such factors as ability, age, and sex. The

---

[8]John Stull, "Implications of the Stull Bill" (paper presented at the Conference of the Association of California School Administrators, Pasadena, Ca., September, 1975).

[9]Gary Saretzky, "The Strangely Significant Case of Peter Doe," Phi Delta Kappan, 54 (May, 1973), pp. 589-92.

[10]Shirley Boes Neill, "The Competency Movement," Critical Issues Series, American Association of School Administrators, Arlington, 1978 (monograph).

[11]James C. Enochs, "Modesto, California: A Return to the Four R's," Phi Delta Kappan, 59 (May, 1978), pp. 609-10.

objectives that are to be mastered usually center around the degree to which learning is assumed to be qualitative rather than quantitative. Assessing objectives this way would presuppose some kind of measurement. The "if" and "how" of assessment has often split administrators along ideological lines.[12]

The argument does not generally focus on specific, observable, behavioral detail, nor on aspects of either learning, objectives, or learner characteristics. Rather, proponents and opponents of assessment of learning might be broadly divided into two philosophical camps that are distinguishable along certain presuppositions. One of them is the use and/or misuse of academic feedback.[13]

Humanistically oriented educators tend to ascribe to a belief system that holds that learning is best inductively taught, and that the objectives of schooling should not be made subject to imperfect psychometric evaluation.[14] Behavioristically oriented

---

[12] Donald W. Robinson, Book Review of Shirley Boes Neill's The Competency Movement, Phi Delta Kappan, 59 (May, 1978), p. 639.

[13] Arthur E. Wise, "Minimum Competency Testing: Another Case of Hyper-Rationalization," Phi Delta Kappan, 59 (May, 1978), pp. 596-98.

[14] John Holt, "I Oppose Testing, Marking, and Grading," Today's Education, 60 (March, 1971), pp. 28-31.

educators see learning as being subject to precise behavioral laws. Educational measurement is seen as guiding future behavior, and as a means of assessing whether learning objectives have been achieved. Understanding and applying the use of feedback is crucial, according to behaviorist thinking.[15] Both schools of thought tend to see feedback in absolute terms--as something that "should" be given or withheld.

American education has had a history of swings between these two poles.[16] While the 1960's were noted for their humanistic orientation, the latter part of the 1970's has seen a swing back to more impartial, output-oriented teaching. The accountability movement with its stress on earned rewards and systematic feedback is evidence of this particular public demand. "Back to Basics" is in part a strong demand for grading. For too long, say noted writers such as Ebel,[17] teaching has been seen ideally but erroneously as an "art"--a practice guided by intuition, with the teacher standing beside or behind self-generated activities, dispensing generous praise regardless of the merits of the product that came forth. Since random

---

[15] Robert L. Ebel, "Educational Tests: Valid? Biased? Useful?" Phi Delta Kappan, 56 (October, 1975), pp. 83-88.

[16] George F. Kneller, Foundations of Education (New York: John Wiley and Sons, Inc., 1971), pp. 231-51.

[17] Ebel, op. cit., p. 547.

activity per se and not the actual quality and/or quantity of output was the criterion of praise, scholastic scores would naturally falter. To counter this argument, a number of writers of opposing convictions have repeatedly stressed that discovery learning carries its own intrinsic reward, regardless of whether or not test scores actually verify gains.[18] Much of the criticism for or against particular modes of teaching comes in the guise of words laden with emotion, used not so much to convey accurate meaning as to discredit or honor a particular belief. Input control and output efficiency, Foshay[19] points out repeatedly, is made subservient to ideological convictions. Opinion and not data, he stresses, is the basis for running the schools.

The controversy surrounding grading, specifically, is colored by strong a priori ideological investments. Proponents of grading have held that motivation to learn cannot be taken for granted without clear knowledge as to results,[20] while opponents of grading

---

[18] William Glasser, Schools Without Failure (New York: Harper and Row, 1969), pp. 69-88.

[19] A. Foshay, "Sources of School Practice," The Elementary School in the United States, ed. J. Goodlad and H. Shane: 72nd Yearbook of the National Society for the Study of Education, Part II (Chicago: University of Chicago Press, 1973).

[20] Ebel, loc. cit.

have argued that there is an inherent injustice in grading in that the brightest children, who need rewards least, will get them disproportionately, while weaker children are systematically short-changed.[21] This argument is especially potent regarding minority children whose output is often weaker than estimated ability would warrant.[22] Justice, say critics of systematic rewards, must be served by withholding discriminate feedback, or by giving equal feedback for unequal products, for not doing so means weakening a youngster's ego strength.[23] Proponents assert that grading is justified "merit pay" in that scholastic output is objectively and consistently rewarded.[24] Belief in the relative merits and/or "misfirings" of grading has historical roots. It seems less clear, however, whether the argument is resting on factual premises.

One might argue that the controversy regarding whether or not to use feedback in the instructional process has obscured

---

[21] Robert L. Green, "Tips on Educational Testing: What Teachers and Parents Should Know," Phi Delta Kappan, 56 (October, 1975), pp. 89-93.

[22] Minerva Mendoza Friedmann, "Spanish-Bilingual Students and Intelligence Testing," Thrust, 3 (November, 1973), pp. 20-23.

[23] Clair C. Coons, "Non-Promotion: A Dead End Road," Phi Delta Kappan, 58 (May, 1977), pp. 701-2.

[24] Robert L. Ebel, "The Case for Minimum Competency Testing," Phi Delta Kappan, 59 (April, 1978), p. 547.

certain intriguing possibilities. Feedback may affect different students differently. For example, feedback may work well for slow children but not so well for bright, or vice versa. It may be more appropriate for rote learning than for more sophisticated, higher-order learning. It may work differently for boys than girls, given our cultural milieu that assigns differential sexual roles and concomitant behavior very early.

While grading may be thought of as the most common systematic feedback or "pay-off" in direct relation to a youngster's academic output, there are other areas of concern to school administration where clear academic feedback may be given or withheld--for example, in the use or rejection of disguising practices of grouping for instructional purposes, in decisions regarding vertical versus horizontal class organization, in "courtesy promotion" where children are kept with age-equal rather than scholastic-equal peers, in the sometimes widespread practices of "grade inflation," in the installing or banning of forms of symbolic merit systems such as honor passes, honor diplomas, or badges, or in the more mundane everyday class practices such as uncorrected or self-corrected papers. As stated before, informational feedback for the sake of instructional improvement seems to be given or withheld according to prevailing ideological beliefs. A finer distinction is seldomly made that would tell how different learners might react to feedback in different kinds of learning tasks, or how feedback may affect more complex, hierarchical learning, as opposed to simple recall.

One might safely project that the accountability movement will intensify the search for sound instructional practices as verified by objective results. Already, a new thrust in school administration seems to have come to the fore. Some writers forecast that administrators, in trying to engineer for results, will come to see themselves increasingly as detached educational managers. A recent publication pertaining to educational change anticipates that administrative practice will impose the methodological discipline other fields demand in their search for better solutions to plaguing problems. To quote these authors, in part:

> ...Men must try to plan their changing futures, and this necessity is seen to be determined by cultural conditions, not primarily by the ideology men happen to hold...This (helps) to account for the shift of many questions about planned change from an ideological to a technical form.[25]

Theoretical Aspects

Theoretical support for the systematic use of feedback techniques can be drawn from respected psychological quarters.

---

[25]Warren G. Bennis, Kenneth D. Benne, Robert Chin, and Kenneth E. Corey, The Planning of Change (New York: Holt, Rinehart, and Winston, 1976), p. 16.

Skinner[26] has repeatedly emphasized that educators cannot afford to leave the acquisition of systematic knowledge to accident, chance, and whim. Refinement of skills occurs for a reason. Specifically, students toil persistently hard for the feedback from a teacher they admire. If such teacher withholds qualitative feedback, he weakens a child's knowledge potential. According to Skinner, a teacher blithely dispatching indiscriminate feedback puts sand into the delicate psychological mechanisms subject to behavioral laws. In contrast, a teacher unafraid to use his shaping skills is not, in Skinner's words, "... increasing the extent of control under which people live. He simply improves on the kinds of controls that are ineffective and troublesome."[27] Skinner, in his empirical work, has repeatedly demonstrated that feedback does work. But does it always work for children of different entry characteristics, such as ability, age, and sex? To highlight a point made earlier, one might argue that bright children, more accustomed to academic success, will react differently to the influence of feedback than slow children. Similarly, developmental milestones may determine that young children will react differently to feedback than older children. Boys may be more

---

[26]"Skinner Testifies on Education," Education Daily (Washington: Capitol Publications, 32, May, 1973).

[27]Ibid., p. 630.

accustomed to objective rewards, while girls may not need as much feedback, or vice versa.

The need for a more detailed scientific experimental foundation regarding feedback properties has also been stressed by Whittock[28] and Cronbach and Suppes.[29] "Deliberate inquiry," say these authors, "must be an essential component of (instructional) development."[30] Other writers have pointed out that a priori conjecture regarding the merits and/or demerits of feedback properties not only allows unexamined practices to continue, but also prevents a number of worthwhile experiments from being conducted.[31] Most administrators, Wilson and Schmits[32] point out, accept the premise that children have different learning capacities. Therefore, differences in speed of learning and in breadth of learning seem plausible. Mastery learning, a relatively new set of accountability

---

[28] M. C. Whittock, "Product-Oriented Research," The Educational Forum, 31 (May, 1967), pp. 145-50.

[29] L. J. Cronbach and P. Suppes, eds., Research for Tomorrow's Schools: Disciplined Inquiry for Education, Report of the Committee on Educational Research of the National Academy of Education (London: MacMillan, Callien Macmillan, Limited, 1969).

[30] Ibid, p. 43.

[31] W. James Popham, Educational Evaluation (Englewood Cliffs: Prentice Hall, Inc., 1975), pp. 1-17.

[32] Barry J. Wilson and Donald W. Schmits, "What's New in Ability Grouping?" Phi Delta Kappan, 59 (April, 1978), p. 535.

strategies, carries a strong component of feedback. McKeachie[33] points out that such feedback, even if given in an informative way, is insufficient for optimum learning. A prescription has to be added that tells how to correct unsatisfactory results. Bloom's[34] term for this technique is "feedback correction procedures"--providing information on how each student's learning is changing as a result of the unit's initial instruction. Mastery learning strategies utilizing feedback techniques do not eliminate the effects of ability differences, but they can minimize them substantially if properly used. In one experiment, at least, it appeared that those weakest in cognitive entry characteristics learned more like their stronger peers once feedback was made part of the mastery strategy design. In other words, the distance between the strongest and weakest was actually narrowed, not broadened, as claimed by opponents of feedback.

---

[33]W. R. McKeachie, "The Decline and Fall of the Laws of Learning," Educational Researcher, 3 (July, 1974), pp. 7-11.

[34]B. S. Bloom, "Mastery Learning and its Implication for Curriculum Development," Confronting Curriculum Reform, ed. E. W. Eisner (Boston: Little, Brown, 1971).

Several authors have focused on type of feedback,[35] while others have studied the dosage.[36] Still others have argued on ideological grounds that a student may <u>want</u> to modify his progress on the basis of his own successes and failures, but that the subsequent test is often deliberately withheld from the child.[37] These writers argue that a student likes to have his answers evaluated. If he has replied correctly, he appreciates recognition of this fact. If incorrectly, he welcomes a lead toward the answer he has missed. An instant appraisal of a child's reply, together with the instruction he may receive if he is incorrect, may give the student an incentive to continue with the task.

Summarily, the literature shows that there are strong arguments suggesting the merits and demerits of feedback as a means of instructional quality control. Lacking, however, is what Jerome S. Bruner calls "practice translation." Says Bruner: "Let us

---

[35] J. Farmer, G. Lachter, J. Blaustein, and D. Cole, "The Role of Proctoring in Personalized Instruction," <u>Journal of Applied Behavioral Analysis</u>, 5 (December, 1972), pp. 401-4.

[36] J. H. Block and M. Tierney, "An Exploration of Two Correctional Procedures Used in Mastery Learning Approaches to Instruction," <u>Journal of Educational Psychology</u>, 66 (November, 1974), pp. 962-67.

[37] M. Scriven, <u>Problems and Prospects for Individualized Education</u> (Berkeley: McCutchan, 1975).

begin with a concrete psychology that occupies itself with wily strategies for learning specific things. . . "[38]

THE PROBLEM

Statement of the Problem

First, what is the evidence regarding learning as a function of informational feedback in global, Gestalt-type task mastery where feedback, given in response to lower order learning, may aid the mastery of higher order tasks?

Second, what evidence exists that, given informational feedback, students will suffer depletion of motivation as a result, as inferred by a decreased quantity and quality of learning?

Third, do entry variables such as ability, age, and sex interact with feedback in a significant way?

Purpose of the Study

The purpose of this study was to investigate the impact of systematic informational feedback on learning in a controlled laboratory setting, using the experimental controls inherent in identical twins. One of the most precise psychometric instruments to

---

[38] Jerome S. Bruner, On Teaching Teachers, "Current Issues in Higher Education," Proceedings of the 19th Annual National Conference on Higher Education, Association for Higher Education (Washington, D. C., 1964), p. 98.

date, the <u>Wechsler Intelligence Scale for Children/Revised,</u>* was used in an unconventional manner--<u>not</u> as a measure of IQ primarily but as a series of hierarchical learning tasks subject to concise experimental manipulation.

Possible differences with regard to different forms of learning, as embodied in the scales of the instrument, were investigated. Possible differences with regard to ability, age, and sex were also examined.

HYPOTHESES

Ho_1: There is no difference between the experimental and control groups on the following WISC-R scales: Information, Similarities, Arithmetic, Vocabulary, Comprehension, Picture Completion, Picture Arrangement, Block Design, Object Assembly, and Coding.

Ho_2: There is no difference between the experimental and control groups on the verbal and non-verbal scales.

Ho_3: There is no difference between the experimental and control groups on the Full Scale Score.

Ho_4: The effects of feedback are constant with respect to ability.

---

* Abbreviated to WISC-R in the following pages.

Ho_5: The effects of feedback are constant with respect to age.

Ho_6: The effects of feedback are constant with respect to sex.

METHODS AND PROCEDURES

Subjects in this study consisted of 30 sets of identical twins from ages 6 to 16. Since twinning occurs randomly in all races and across broad sociological and demographic lines and follows no discernible genetic pattern,[39] less attention could be given to random sampling procedures, and more effort was spent minimizing attrition.

Proximity and willingness to participate dictated, in part, the choice of the twins who were part of this study. All came from the vicinity of Stockton, California. A detailed description of methods of sampling is given in Chapter 3.

Each pair of twins was randomly assigned into an experimental and control group. The control group took the WISC-R according to standardized instructions. The WISC-R Manual prohibits feedback as to the quality and/or quantity of the subject's response; therefore, the control group was "left in the dark" as to how well or how poorly they did in solving individual items.

---

[39] Encyclopedia Britannica, Volume 15 (Chicago, 1971), p. 985.

The experimental group was told at the outset that items would be progressively more difficult, and that there was no way all items could be solved. They were told they would be informed whether items were solved "right" or "wrong." In the case of failed items, precise correction would be given.

Test instructions specify that approximately one-third of the items had to be failed before a "ceiling" was reached and testing could be discontinued. These were the items left uncorrected for the control group, while the experimental group was given feedback as to what a "perfect" answer would have been. It was theorized that this correction, given in response to lower order items, would result in a higher performance on subsequently more difficult tasks, with possible differentiation as to ability, age, and sex.

All testing was done outside of school. All children were tested individually. Answers were recorded on special WISC-R protocol sheets and later transferred to key punch cards for statistical computer analysis.

ASSUMPTIONS

Globally, this study was based on the assumption that human behavior is orderly, lawful, and predictable. Learning, as evidenced by scholastic output in school, was seen as one aspect of behavior that might be improved by systematically varying and controlling informational input.

Specifically, this study was predicated on two fundamental assumptions. The primary assumption was that the WISC-R scales were indeed hierarchical. The second assumption was that the task of solving individual items was representative of global cognitive abilities.

LIMITATIONS AND DELIMITATIONS

The extensive individual testing required by this study was very expensive, time-consuming, and cumbersome. Therefore, only one tester (the investigator) was involved. This may have introduced an unconscious degree of examiner bias. However, every effort was made to keep bias to a minimum.

A further constraint was dictated by the unassessed demographic characteristics of this particular twin sample. No attempt was made to obtain a cross-section of subjects on socio-economic parameters such as parental income, education, or race.

Finally, practical constraints dictated by the choice of the instrument demanded that this study be limited to ages 6 through 16. Any generalizations drawn from this study pertain only to this age group.

DEFINITION OF TERMS

Accountability: public demand on school administrators to provide measurable scholastic results in efficient relation to instructional input.

Behaviorism: a system of beliefs stressing that behavioral improvement is possible through the systematic use of feedback techniques.

Educational quality control: the administrative effort to channel resources effectively and efficiently.

Efficiency: the amount of measurable academic output in relation to a specified amount of instructional input.

Feedback: precise information as to the correctness or incorrectness of a learner's response, using the given criteria of the WISC-R Manual.

Grading: the quantification of behavioral data by teachers through the use of letters or numbers.

Input: instructional means, tactics, or resources used to increase measurable scholastic results.

Humanism: a system of beliefs that stresses, in part, that intrinsic educational values are presently inaccessible to precise, impartial quantification.

Output: measurable learning in the form of psychometric scores.

Twin: a member of a set of identical twins, reared together.

SIGNIFICANCE OF THE STUDY

This study demonstrated the effects of informational feedback in ten specific, discrete hierarchical learning tasks in light of learner variables such as ability, age, and sex. The investigation attempted to clarify important questions regarding the educational costs and/or benefits accrueing to instruction through the deliberate, conscious, and specialized use of feedback. Hereby, the study demonstrated the linkage of feedback properties to educational/ instructional quality control. More broadly still, the study attempted to shed light on the "ideal" form of education as loosely held by traditional beliefs, versus the pragmatic conditions under which schools operate.

The results, it is hoped, will aid administrators in making informed decisions regarding instructional policies in response to public demand for accountability. These decisions are likely to affect the practices of grading, grouping for instructional purposes, vertical versus horizontal class organization, the possible use of symbolic merit systems, and others.

SUMMARY

Rather than an administrative "either-or" approach regarding instructional feedback, an alternative approach was suggested pertaining to the judicious use of feedback in particular light of certain characteristics of learning, and in light of learner variables such as ability, age, and sex. A learning setting embodying progressively more difficult tasks subject to precise experimental control, with feedback as the experimental variable and various psychometric results as the dependent variables, was seen as providing useful and substantive answers to the hypotheses posed in this chapter.

Chapter 2

REVIEW OF THE LITERATURE

INTRODUCTION

Proponents of feedback principles stress that an environment highly responsive to a behavior in an objective, discriminating way will facilitate the re-occurrence of that behavior and/or behaviors that are similar or close.[1] Applied to classroom learning, an environment highly responsive to scholastic output in an objective, discriminating way is believed to facilitate achievement.[2] It was stressed in Chapter 1 that educational leaders seem to have given fluctuating attention to this contingency of learning, depending upon prevailing ideological beliefs. However, giving individuals (or groups) practice in what they are expected to do, and informing them clearly and consistently of their successes or failures, is an accepted principle in certain sub-areas of learning management. Two major classroom applications of informational feedback will be

---

[1]B. F. Skinner, Science and Human Behavior (New York: MacMillan, 1953), p. 18.

[2]B. F. Skinner, The Technology of Teaching (New York: Appleton-Century-Crofts, 1968), p. 26.

reviewed: 1) operant conditioning as found in programmed instruction[3] and computer-assisted instruction,[4] and 2) mastery learning as a currently popular instructional strategy.[5]

Both operant conditioning and mastery learning as classroom strategies are fairly new in origin, having grown out of dissatisfaction with open classroom practices and their concomitant withholding of feedback.[6] As applied to classroom management, both are seen as strategies that aid what is known as "cognitive learning." In both, feedback is viewed as an important and recognized ingredient of educational quality control. It is of interest to note that feedback as an essential component in educational management seems to be better recognized in practices pertaining to non-cognitive learning, especially psychomotor learning. For example, Ammons'[7] review and

---

[3]A. A. Lumsdaine and R. Glaser, eds., Teaching Machines and Programmed Learning: A Source Book (NEA, 1960).

[4]A. A. Lumsdaine, "Instruments and Media of Instruction," Handbook of Research on Teaching, ed. N. L. Gage (Chicago: Rand McNally, 1963), pp. 583-682.

[5]James H. Block and Robert B. Burns, "Mastery Learning," Review of Research in Education, ed. Lee S. Shulman (Itaska: Peacock Publishers, Inc., 1976), pp. 3-41.

[6]B. F. Skinner, "The Free and Happy Student," Phi Delta Kappan, 55 (September, 1973), pp. 13-16.

[7]R. B. Ammons, "Effects of Knowledge of Performance: A Survey and Tentative Theoretical Formulation," Journal of General Psychology, 54 (1956), pp. 279-299.

theoretical formulation pertaining to knowledge of results shows that most experimental studies from which he drew his generalizations were in the area of psychomotor development. In the area of visceral learning,[8] performance is also seen as directly proportional to the availability of feedback as a means of behavioral quality control. Subsequent pages will give a cursory look at the role of feedback in non-cognitive endeavors as an introduction to feedback in cognitive learning management.

Selected studies focusing on the learner variables of ability, age, and sex in relation to feedback will also be cited. These studies will serve to illustrate that learning--however well controlled experimentally it may be--does not happen in isolation. Developmental psychologists have repeatedly stressed that learning is vastly different for young versus older children,[9] somewhat different for boys than girls,[10] and appears to be different, qualitatively, for

---

[8]N. E. Miller, "Extending the Domain of Learning," Science, 152 (1966), p. 676.

[9]J. Piaget, The Science of Education and the Psychology of the Child (New York: Grossman, 1970).

[10]C. Broderick and J. Bernard, eds., The Individual, Sex, and Society: A SIECUS Handbook for Teachers and Counselors (Baltimore: The Johns Hopkins Press, 1969).

bright versus slower children.[11] It is not germane to the purpose of this study to sort out the numerous findings pertaining to ability, age, and sex at this time, or to cite exhaustive evidence that differences in learning appear to be the result of complex transactions between genetic and constitutional factors and environmental influences. A representative sample of studies illustrating the complexities in children with respect to ability, age, and sex in light of feedback principles, both cross-sectionally and longitudinally, will serve in part to warrant the focus on the problems addressed in this study.

Finally, since informational feedback was investigated empirically in a hierarchically-structured setting, it seemed appropriate to review the major theories, studies, and findings pertaining to hierarchical learning, especially those by Bruner,[12] Ausubel,[13] and Gagne.[14]

---

[11] J. W. Getzels and J. T. Dillon, "The Nature of Giftedness and the Education of the Gifted," Second Handbook of Research on Teaching, ed. Robert W. Travers (Chicago: Rand McNally and Co., 1973), pp. 689-731.

[12] J. S. Bruner, R. R. Olver, and P. M. Greenfield, eds., Studies in Cognitive Growth (New York: Wiley, 1966).

[13] D. P. Ausubel, Educational Psychology: A Cognitive View (New York: Holt, Rinehart and Winston, 1968).

[14] Robert M. Gagne, The Conditions of Learning (New York: Rinehart and Winston, 1965).

FEEDBACK PRINCIPLES IN NON-CLASSROOM ACTIVITIES

Motor Learning

With the stimulus provided by Ammons'[15] review and theoretical formulation, physical educators have made the role of adequate feedback paramount in their instructional strategies. Bilodeau and Bilodeau[16] stress that the motivational and corrective influences of feedback are the most important variables controlling motor performance. Methods of augmenting feedback are continually being refined. Many attempts have been made to increase the potency of feedback by providing a model of an "ideal" performance to steer subsequent learning efficiently.[17] There are numerous studies exploring the merits of kinesthetic feedback upon learning,[18] all suggesting that physical learning is either enhanced or impaired

---

[15] R. B. Ammons, "Effects of Knowledge of Performance: A Survey and Tentative Theoretical Formulation," Journal of General Psychology, 54 (1956), pp. 279-299.

[16] E. A. Bilodeau and I. M. Bilodeau, "Motor Skills Learning," Annual Review of Psychology, 12 (1961), pp. 243-280.

[17] M. L. Howell, "Use of Forcetime Graphs for Performance Analysis in Facilitating Motor Learning," Research Quarterly, 27 (1956), pp. 12-22.

[18] John E. Nixon and Lawrence F. Locke, Second Handbook of Research in Teaching, ed. Robert M. W. Travers, American Educational Research Association (Chicago: Rand McNally, 1973), pp. 1210-42.

according to the availability or non-availability of feedback. The learner's attention has to be drawn to the quality of his own performance. To the extent that this information is withheld, performance will invariably suffer.[19] Malina[20] states that manipulation of feedback directly influences motor learning, especially visual learning in the early stages. If feedback is provided early, both speed and accuracy of performance are improved.[21] These writers make the distinction between two categories of feedback--feedback in the form of knowledge of performance, and feedback in the form of knowledge of results.

Compared to most research pertaining to classroom learning, these studies seem relatively unambiguous and clear-cut. It is easier to measure motor learning than invisible "mental processes" such as occur in cognitive learning. Most studies in motor learning have clear designs, adequate controls, and acceptable criteria. Limitations, however, might be lack of generalizability due

---

[19] J. D. Lawther, The Learning of Physical Skills (Englewood Cliffs, N. J.: Prentice-Hall, 1968), p. 8.

[20] R. M. Malina, "Effects of Varied Information Feedback Practice Conditions on Throwing Speed and Accuracy," Research Quarterly, 40 (1969), pp. 135-45.

[21] J. Annett and H. Kay, "Knowledge of Results and 'Skilled Performance'," Occupational Psychology, 31 (1957), pp. 69-79.

to investigations on very narrow aspects of motor learning, such as studies on knob turning or lever positioning.

Visceral Learning

Evidence of the controlling facets of feedback can also be found in the realm of "unconscious" or "automatic" learning.[22] Visceral responses such as galvanic skin responses, heart rate, heart rhythm, salivation, blood pressure, etc.--once thought to be "spontaneous" and not directly amenable to learning--can be modified, monitored, and controlled if information regarding their magnitude is "at the subject's fingertips." Information provided by a device that translates these biological functions onto a visual chart for purposes of quick assessment is called biofeedback. Biofeedback principles have not yet found their way into the classroom, although some preliminary suggestions have been made.[23] The most obviously beneficial application might be the use of biofeedback devices for purposes of reduction of anxiety and stress. Readers are referred to

---

[22] N. E. Miller, "Learning of Visceral and Glandular Responses," Science, 163 (1969), pp. 434-45.

[23] N. E. Miller, "Experiments Relevant to Learning Theory and Psychopathology," Psychopathology Today: Experimentation, Theory, and Research, ed. W. S. Sahakian (Itaska, Ill.: Peacock, 1970), pp. 148-66.

an excellent summary article by Miller,[24] a pioneer authority in biofeedback, listed in the 1978 edition of the Annual Review of Psychology.

FEEDBACK PRINCIPLES IN CLASSROOM ACTIVITIES

Programmed Instruction and Computer-Assisted Instruction

Operant conditioning basically holds that a response must be emitted and reinforced by feedback in order to come under stimulus control.[25] Therefore, two important attributes of operant conditioning are efficient provisions for learner response and immediate feedback as to the quality and magnitude of the response. The degree to which these provisions are presently accommodated in the classroom are left largely to the discretion of the educational staff, and to the individual teacher's awareness of the importance of making these provisions. This awareness at present seems weak, according to B. F. Skinner,[26] one of the most outspoken proponents of

---

[24]Neal E. Miller, "Biofeedback and Visceral Learning," Annual Review of Psychology, 29 (1978), pp. 373-404.

[25]Graham Nuthall and Ivan Snook, "Contemporary Models of Teaching," Second Handbook of Research in Education, ed. Robert M. Travers (Chicago: Rand McNally and Co.), pp. 54-9.

[26]"Skinner Testifies on Education," Education Daily (Washington: Capitol Publications, May 1973).

operant conditioning principles. One controversial application of Skinner's writings has been the use of programmed instruction and its variation, computer-assisted instruction.[27]

Thorndike[28, 29] was one of the first researchers to recognize the potential value of knowledge of results in classroom instruction. But it was Pressey[30] who, beginning in the second decade of this century, first developed devices that automated responses and feedback. His machines provided immediate qualitative evaluation of subjects taking a test. His intent was to supplement classroom instructions by providing the learner with rapid knowledge of results for corrective guiding purposes. Pressey's experiment was not so much an empirical verification of a principle as it was a physical labor-saving device with built-in feedback properties--the forerunner of the modern computer. Many other educational "technicians" have built upon the precedents set by

---

[27]Nuthall and Snook, op. cit., p. 58.

[28]Edward L. Thorndike, The Psychology of Learning (New York: Columbia University, 1913).

[29]Edward L. Thorndike, "The Law of Effect," American Journal of Psychology, 39 (1927), pp. 212-22.

[30]S. L. Pressey, "A Simple Apparatus Which Gives Tests and Scores--and Teaches," School and Society, 23 (1926), pp. 373-6.

Pressey.[31] All reported that immediate feedback was an efficient condition for learning. These experiments were all empirical in nature and added to the cumulative evidence that knowledge of results aided learning.

B. F. Skinner, in his now classic "The Science of Learning and the Art of Teaching," introduced operant conditioning principles to broad classroom application.[32] His work was quickly replicated and refined by Lumsdaine and Glaser,[33] Stolurow,[34] and Dale.[35]

Lumsdaine[36] provided the first broad coverage of the literature of knowledge of results, summarizing and ordering the work done up to that time. Since then, several excellent reviews have

---

[31]John E. Nixon and Lawrence F. Locke, op. cit., pp. 619-21.

[32]B. F. Skinner, "The Science of Learning and the Art of Teaching," Harvard Educational Review, 24 (1954), pp. 86-97.

[33]A. A. Lumsdaine and R. Glaser, eds., Teaching Machines and Programmed Learning: A Source Book (National Education Association, 1960).

[34]L. M. Stolurow, "Teaching by Machine," Cooperative Research Monographs, 6 (1961), pp. 14-20.

[35]Edgar Dale, "Historical Setting of Programmed Instruction," Programmed Instruction, ed. Phil C. Lange (University of Chicago Press, 66th Yearbook, NSSE, 1967).

[36]A. A. Lumsdaine and R. Glaser, "Instruments and Media of Instruction," Handbook of Research on Teaching, ed. N. L. Gage (Rand McNally, 1963), pp. 583-682.

appeared, including those by Schramm,[37] Holland,[38] May,[39] Popham,[40] and Gagne and Rohwer.[41] The following is a summary of the previous studies and the vast array of converging evidence regarding the merits of knowledge of results, as summarized in part by McKeachie.[42]

In general, knowledge of results matters little if it follows correct responses that are known to be correct, but matters greatly in the acquisition of learning following either incorrect responses or correct responses where the learner was in doubt as to the proximity to an "ideal" response. The value of feedback seems to be stronger in the case of incorrect responses. Such feedback

---

[37]W. L. Schramm, The Research on Programmed Instruction (Washington, D. C.: U. S. Government Printing Office, 1964), pp. 88-94.

[38]J. G. Holland, "Research on Programing Variables," Teaching Machines and Programmed Learning, II: Data and Directions, ed. R. Glaser (Washington, D. C.: National Education Association, 1965), pp. 66-117.

[39]M. A. May, The Role of Student Response in Learning from the New Educational Media (Washington, D. C.: U. S. Dept. of Health, Education and Welfare, Final report, USOE Contract OE-5-16-006, 1966).

[40]W. J. Popham, "Curriculum Materials," Review of Educational Research, 39 (1969), pp. 319-38.

[41]R. M. Gagne and W. D. Rohwer, Jr., "Instructional Psychology," Annual Review of Psychology, 20 (1969), pp. 381-418.

[42]Wilbert J. McKeachie, "Instructional Psychology," Annual Review of Psychology, 25 (1974), pp. 161-66.

permits the learner to correct his mistakes, and lessens the likelihood that wrong responses will be repeated. Furthermore, feedback supplying a "right" answer seems to be better than simply telling the learner he was wrong. In addition, learning appears to be a direct function of the speed, magnitude, and quality of the information given. However, there are at least two exceptions: Sassenrath and Yonge[43] showed that delayed feedback strengthened recall, and More[44] found that intermediate feedback delay was better than either immediate feedback or feedback coming after an extensive delay. Sassenrath[45] states that there is no advantage for immediate over delayed feedback. Sassenrath's study encompassed a large sample (N = 311) and a complex design, but "feedback" seemed to have been arbitrarily defined, and the findings, though significant, were tenuous. Sturges[46] speculated that the advantage of delayed feedback

---

[43] J. M. Sassenrath and G. D. Yonge, "Effects of Delayed Information Feedback and Feedback Cues in Learning on Delayed Retention," Journal of Educational Psychology, 60 (1969), pp. 174-77.

[44] A. J. More, "Delay of Feedback and the Acquisition and Retention of Verbal Materials in the Classroom," Journal of Educational Psychology, 60 (1969), pp. 339-42.

[45] J. M. Sassenrath, "Effects of Delay of Feedback and Length of Postfeedback Interval on Retention of Prose Material," Psychology in Schools, 9 (1972), pp. 194-97.

[46] P. T. Sturges, "Information Delay and Retention: Effect of Information in Feedback and Tests," Journal of Educational Psychology, 63 (1972), pp. 32-43.

may be due to the subjects' exploring the organization of material before feedback. Like Sturges, other researchers[47] suspected the reason for this finding to be the subjects' lackadaisical attitude in situations where it was known that feedback was forthcoming, thus encouraging inattentive and careless behavior on the learner's part. Kulhavy and Anderson,[48] studying delayed feedback, suggest again the importance of the actual magnitude of the information provided. Other important factors are that the feedback be placed strategically and that the subject still remembers what he is being reinforced for. The last four studies have a strongly inferential component in that the authors seemed to engage in inferring motivational attributes not warranted by the empirical facts of the studies.

Olson,[49] in a longitudinal study on retention of learning, found that feedback lost its power over time, as did Oner,[50] by

---

[47] R. C. Anderson, R. W. Kulhavy, and T. Andre, "Feedback Procedures in Programmed Instruction," Journal of Educational Psychology, 62 (1971), pp. 148-56.

[48] R. W. Kulhavy and R. C. Anderson, "Delay-retention Effect with Multiple-choice Tests," Journal of Educational Psychology, 63 (1972), pp. 505-12.

[49] G. H. Olson, "A Multivariate Examination of the Effects of Behavioral Objectives, Knowledge of Results, and the Assignment of Grades on the Facilitation of Classroom Learning," PhD thesis, Dissertation Abstracts International, 32 (1972), pp. 6214-15.

[50] N. P. Oner, "Impact of Teacher Behavior and Teaching Technique on Learning by Anxious Children," Dissertation Abstracts International, 32 (1972), p. 6215.

studying feedback in programmed math lessons. Moreover, feedback did not interact with anxiety in determining performance. The crucial factor seemed to be the amount of information given by knowledge of results, and the learners' adequate strategies in making use of the feedback given. In the case of programmed learning, incremental steps are very small. Therefore, knowledge of results conveys little useful information. If tasks become more difficult, one might speculate that informational feedback might gain proportionately in corrective power and control.

Empirical studies by Meyer,[51] Holland,[52] and Porter[53] have demonstrated that devices embodying feedback principles can teach. These studies seem carefully designed and adequately controlled. However, several questions remain unanswered. For example, the Holland study used a very selective sample (Harvard and Radcliffe students). Porter's study, under sponsorship by USOE, was exploratory in nature, and not expected to give definite answers.

---

[51] Susan R. Meyer, "A Program in Elementary Arithmetic: Present and Future," Automatic Teaching: The State of the Art, ed. E. Galanter (Wiley, 1959), pp. 83-4.

[52] J. G. Holland, "A Teaching Machine Program in Psychology," Automatic Teaching: The State of the Art, ed. E. Galanter (Wiley, 1959), pp. 69-82.

[53] D. Porter, "Some Effects of Year-long Teaching Machine Instruction," Automatic Teaching: The State of the Art, ed. E. Galanter (Wiley, 1959), pp. 85-90.

Costin,[54] in a review of lecture versus programmed teaching, concluded that programmed study had an advantage in promoting acquisition of learning. Similar findings were reported by Beard and Bligh's[55] comprehensive review.

There have been variations of programmed instruction expanding Skinner's original linear strategies to what has come to be known as "branching."[56] Unlike linear programming, this variation takes into account the quality and proximity of the learner's response as compared to an "ideal" response. Some of these strategies have been described by Crowder as "intrinsic programming"[57] and by Stolurow as "ideomorphic programming."[58]

---

[54]F. Costin, "Lecturing Versus Other Methods of Teaching: A Review of Research," British Journal of Educational Technology, 3 (1972), pp. 4-30.

[55]R. M. Beard and D. A. Bligh, Research into Teaching Methods in Higher Education, Higher Education Monographs (London: Sociological Research in Higher Education, 3rd ed., 1972).

[56]N. A. Crowder, "Automatic Tutoring By Means of Intrinsic Programming," Automatic Teaching: The State of the Art, ed. E. Galanter (Wiley, 1959), pp. 109-16.

[57]N. A. Crowder, "Automatic Tutoring by Intrinsic Programming," Teaching Machines and Programmed Learning: A Source Book, ed. A. A. Lumsdaine and R. Glaser (Washington, D. C.: National Education Association, 1960), pp. 286-98.

[58]L. M. Stolurow, "A Model and Cybernetic System for Research on the Teaching-Learning Process," Programmed Learning, 2 (1965-b), pp. 138-57.

Computer-assisted instruction (CAI) is one adaptation of programmed instruction that utilizes both linear and branching strategies. Here is a well-controlled instructional situation whereby the promptness and adequacy of feedback is assured by mechanical means. Fletcher and Atkinson[59] report CAI to be effective, as measured by both criterion and norm-referenced measurements, more so for boys than for girls. Vinsonhaler and Bass[60] reviewed ten major reasonably well-designed studies of the effectiveness of computer-assisted drill, reporting also on the effectiveness of this kind of learning via feedback. Unfortunately, all but two of the studies had inadequate controls.

In general, it may be concluded that research on the various aspects of operant conditioning principles tends to show that students who use it learn quickly and efficiently. Furthermore, highly responsive mechanized environments have increased curious and exploratory behavior[61] in some instances, contrary to popular

---

[59] J. D. Fletcher and R. C. Atkinson, "Evaluation of the Stanford CAI Program in Initial Reading," Journal of Educational Psychology, 63 (1972), pp. 597-602.

[60] J. F. Vinsonhaler and R. K. Bass, Ten Major Studies of the Evaluation of CAI Drill and Practice, Informational Systems Laboratory Report No. 21 (Lansing: Michigan State University, 1971).

[61] H. Fowler, Curiosity and Exploratory Behavior (New York: Macmillan, 1965).

belief that a "discriminating" or "manipulative" environment causes stereotyped, non-creative behavior. Stolurow, speaking of feedback as utilized in programmed instruction, puts the matter in broader perspective when he says:

> The concepts originally systematized in the development of programmed instruction are beginning to influence the management of the classroom. . . this trend, particularly in the use of (systematic) feedback, will probably increase. The decision to use these concepts in and for education are not simple and easy, but any doubts about their permanence or effectiveness would have to stem from prejudice or ignorance. The only course of action for areas of application is how to use (feedback) most effectively, and the only course for research is how to improve upon what we now know so that we can begin to understand teaching and learning as they take place in schools, universities, and training establishments. . .[62]

Mastery Learning

More widely accepted than straight operant conditioning strategies as manifested in programmed instruction and computer-assisted instruction is a package of teaching strategies known as "mastery learning," where feedback is utilized in part. Underlying mastery learning is an explicit assumption that under appropriate instructional conditions, virtually all students can learn

---

[62] L. M. Stolurow, "Programmed Instruction," Encyclopedia of Educational Research, 4th ed. (1969), p. 1020.

well.[63] Furthermore, there is a built-in pedagogical stipulation that teachers employ strategies so that all students do learn well. According to this school of thought, motivation to learn is a consequence of, not a prerequisite to, efficient learning and teaching.

Five essential components form what has become known as "Bloom's LFM (Learning-for-Mastery) strategies."[64] First, a teacher must personally provide feedback on the learner's particular errors. Second, both formative and summative evaluations are built into the learning units in order to assure that feedback is properly spaced. Third, feedback in the form of correctives is carefully keyed to each item in the unit's formative test. Fourth, in an outset orientation students are given explicit instructions as to how they are expected to learn, and to what level of efficiency they are expected to learn. Finally, grading is a function of mastery, not effort.

A variation of Bloom's Learning-for-Mastery approach is the PSI (Personalized-Systems-of-Instruction) approach.[65] This is a

---

[63] J. H. Block and L. W. Anderson, Mastery Learning in Classroom Instruction (New York: Macmillan, 1975).

[64] Benjamin S. Bloom, "Mastery Learning and Its Implications for Curriculum Development," Confronting Curriculum Reform, ed. E. W. Eisner (Boston: Little, Brown, 1971), p. 14.

[65] James H. Block and Robert B. Burns, "Mastery Learning," Review of Research in Education, ed. Lee S. Shulman (Itaska: F. E. Peacock Publications, 1976), p. 9.

student's individually paced approach to mastery. Some of the features of this variation have been summarized by Hartley[66] as follows.

Immediate knowledge of results concerning the appropriateness or correctness of a response is fed to the student. Often, student proctors are used to provide immediate feedback to learners regarding the quality and quantity of their learning. As with Learning-for-Mastery, this approach assumes that virtually all students can master a great deal if

> . . . the instruction is approached systematically, if students are helped if and when they have learning difficulty, if they are given sufficient time to achieve mastery, and *if there is some clear criterion on what constitutes mastery*. . .[67] (author's italics)

As to how well these two strategies work, there is as yet sparse evidence since both approaches are new. The evidence that has accumulated, however, points to encouraging results.

---

[66] J. Hartley, "Programmed Instruction 1954-1974: A Review," *Programmed Learning and Educational Technology,* 11 (1974), pp. 278-91.

[67] Benjamin S. Bloom, "An Introduction to Mastery Learning Theory," *Schools, Society, and Mastery Learning,* ed. J. H. Block (New York: Holt, Rinehart and Winston, 1974), p. 6.

Vanjo and Nicholson[68] investigated integrative learning skills by means of a hybrid LFM/PSI approach, with positive conclusions. Berry,[69] in a similar vein, investigated PSI in an introductory philosophy course and found students using PSI strategies more skillful in language synthesis than students learning conventionally. These studies showed gains, but feedback as the independent variable was not sufficiently isolated to permit strong conclusions as to its merits and usefulness in a particular setting.

Mastery strategies using feedback can also apparently reduce the variability with which students achieve. Arlin[70] found that individual differences in elapsed study time were reduced from 7:1 to 4:1 by feedback methods, and that the mastery-taught students exhibited progressively less variety compared to students learning

---

[68] J. P. Vanjo and S. J. Nicholson, "A Course in Law and Technology," IEEE Transactions on Education, E-18 (1975), pp. 127-31.

[69] G. Berry, "The Keller Method in Introductory Philosophy Courses," PSI: 41 Germinal Papers, ed. J. G. Sherman (Menlo Park: W. A. Benjamin, 1974).

[70] M. N. Arlin, "The Effects of Formative Evaluation on Student Performance," Contemporary Issues in Educational Testing, ed. H. F. Crombag and D. N. DeGruiter (Paris: Mouton, 1974), pp. 67-74.

traditionally. Rosati[71] verified that variability as measured by time required for mastery decreased over time. Lloyd and Knutzen[72] did not come up with conclusive results, however, in a similar study. Davis[73] found that mastery-taught students would study earlier and be more expeditious in their expenditure of time, thus indirectly verifying the hypothesis that motivation does not suffer when impartial feedback is applied. One criticism applied to the foregoing studies is that the term "mastery-taught student" is too broad and all-encompassing, and that no attempt was made to isolate and/or analyze interaction effects.

There is also some evidence that mastery-taught as opposed to non-mastery-taught students exhibit higher-order learning. Only a few of these studies will be cited. Born, Gledhill, and

---

[71] P. Rosati, "A Comparison of the Personalized Systems of Instruction with the Lecture Method in Teaching Elementary Dynamics," Behavior Research and Technology in Higher Education, ed. J. M. Johnson (Springfield: Charles C. Thomas, 1975), pp. 77-83.

[72] K. Lloyd and N. Knutzen, "A Self-paced Programmed Undergraduate Course in the Experimental Analysis of Behavior," Journal of Applied Behavior Analysis, 2 (1969), pp. 125-33.

[73] M. L. Davis, "Mastery Test Proficiency Requirement Affects Mastery Test Performance," Behavior Research and Technology in Higher Education, ed. J. M. Johnston (Springfield, Ill.: Charles C. Thomas, 1975), pp. 120-9.

Davis;[74] Breland and Smith;[75] Cole, Martin, and Vincent;[76] McMichael and Corey;[77] and Sheppard and MacDermot[78] all found evidence of higher order skills mastery after exposure to LFM strategies. These findings, however, were not confirmed by Poggio[79] who studied retention longitudinally and found that higher order learning was lost more easily over time, while lower-order learning

---

[74] D. G. Born, S. M. Gledhill, and M. L. Davis, "Examination Performance in Lecture-Discussion and Personalized Instruction Courses," Journal of Applied Behavioral Analysis, 5 (1972), pp. 33-43.

[75] N. S. Breland and M. P. Smith, "A Comparison of PSI and Traditional Methods of Instruction for Teaching Introduction to Psychology" (paper presented at the National Conference on Personalized Instruction in Higher Education, February, 1974).

[76] C. Cole, S. Martin, and J. Vincent, "A Comparison of Two Teaching Formats at the College Level," Behavior Research and Technology in Higher Education, ed. J. M. Johnston (Springfield, Ill.: Charles C. Thomas, 1975), pp. 45-55.

[77] J. S. McMichael and J. R. Corey, "Contingency Management in an Introductory Psychology Course Produces Better Learning," Journal of Applied Behavioral Analysis, 2 (1969), pp. 79-83.

[78] W. C. Sheppard and H. G. MacDermot, "Design and Evaluation of a Programmed Course in Introductory Psychology," Journal of Applied Behavior Analysis, 3 (1970), pp. 5-11.

[79] J. Poggio, "Long-term Cognitive Retention Resulting from the Mastery Learning Paradigm" (paper presented at the annual meeting of the American Educational Research Association, San Francisco, April, 1976).

remained relatively stable. A similar study by Breland and Smith[80] found similar results. Block and Burns sum up the research by saying:

> In quantitative terms, mastery approaches have usually produced greater student learning than non-mastery approaches, and they have usually produced relatively less variability in this learning. In qualitative terms, mastery approaches have typically helped students acquire higher order learning, though there is some question as to whether this higher order learning has been retained.[81]

In addition, not all students benefit from feedback approaches in the same manner. How they start out initially seems to determine performance. An extensive summary pertaining to student entry characteristics is offered by Block and Burns.[82] Individual ability differences seem to matter the most. While feedback strategies do not eliminate these differences, they seem to minimize

---

[80]N. S. Breland and M. P. Smith, "Cognitive and Affective Outcomes of PSI Mastery Programs as Compared to Traditional Instruction" (paper presented at the annual meeting of the American Educational Research Association, Washington, D. C., March-April, 1975).

[81]Block and Burns, op. cit., p. 25.

[82]Block and Burns, op. cit., pp. 33-36.

their effects.[83, 84] The two keys seen as having this homogenizing effect are clear initial performance demands by the teacher and the expeditious use of feedback correction procedures.

A few studies have concentrated on motivational consequences following the use of objective feedback. Again, to quote Block and Burns:

> ...the mastery approaches have typically elicited more favorable affective responses from students than their non-mastery counterparts, and in some cases significantly more favorable responses...the mastery strategies have had a positive impact on students' interest in and attitudes toward the subject matter learned, self-concept...academic self-confidence, attitudes toward cooperative learning, and attitudes toward instruction.[85]

Block and Burns believe there is sufficient evidence that mastery-taught students have exhibited greater learning, on the average, than their non-mastery counterparts, as measured by both immediate achievement as well as retention over time. On the average, LFM taught students scored five-eighths of a standard deviation better than non-mastery taught students. Likewise,

---

[83] C. K. Burrows and J. R. Okey, "The Effects of a Mastery Learning Strategy on Achievement" (paper presented at the annual meeting of the American Educational Research Association, Washington, D. C., March-April, 1975).

[84] J. R. Nazarro, J. C. Todorov, and J. N. Nazarro, "Student Ability and Individualized Instruction," Journal of College Science Teaching, 2 (1972), pp. 29-30.

[85] Block and Burns, op. cit., p. 26.

LFM/PSI-taught students scored two-thirds of a standard deviation better than non-mastery-taught students, as measured by retention.[86] Similar results were found by Kim,[87] Lee,[88] and Okey.[89] Campbell sums up these studies by stating:

> These are intriguing findings which raise hope that teaching for mastery need not permanently hold back the faster students for the slower, or extend the education of some youngsters until they are oldsters.[90]

It appears that the studies on mastery learning are not directly comparable due to lack of instrument standardization and variations in research design. Often, the populations under study were widely divergent. The most prominent limitation, however, is

---

[86] Block and Burns, op. cit., p. 19.

[87] Y. Kim, "An Application of a New Instructional Model," Research Report No. 8 (Seoul, Korea: Korean Educational Development Institute, 1974).

[88] Y. D. Lee, "Interaction Improvement Studies on the Mastery Learning Project," Final Report on Mastery Learning Program, April-November 1971 (Seoul, Korea: Educational Research Center, Seoul National University, 1971).

[89] J. R. Okey, "Development of Mastery Teaching Materials," Final Evaluation Report, USOE G-74-2990 (Bloomington: Indiana University, August, 1975).

[90] D. T. Campbell and J. C. Stanley, "Experimental and Quasi-experimental Designs for Research on Teaching," Handbook of Research on Teaching, ed. N. L. Gage (Chicago: Rand McNally, 1963), p. 25.

that the term "mastery learning" is composite in nature and encompasses many different variables, of which feedback is only one.

FEEDBACK AND AGE, SEX, AND ABILITY CHARACTERISTICS

There is excessive amount of overlap and duplication in the literature reporting on learner attributes such as ability, age, and sex with respect to aspects of feedback. The one outstanding characteristic is lack of clear focus. For example, sex differences have been investigated with regard to intellectual development, parent, peer, and teacher feedback, and achievement in both young and older children, cross-sectionally and in studies over time. The following studies are offered as representative samples.

A longitudinal study by Kogan and Pankove[91] on the stability of creative ability between grades five and ten found that boys and girls differed increasingly in creative output--boys performing more reliably in group situations where feedback was openly available, while girls seemed to need a more nurturing, individual, attentive approach where feedback was tailored to the needs of each child. A correlation between output and IQ for girls could not be obtained, while for boys at grade ten a positive

---

[91] N. Kogan and E. Pankove, "Creative Ability over a Five-year Span," Child Development, 43 (1972), pp. 427-42.

correlation between IQ and qualitative output was noted. There are some weaknesses in this study. For example, while the pretesting was done individually, the post-testing was, in part, group-administered. This split was made along sex dimensions, for no discernible reasons. The authors conclude: "The findings offered by the present study, though intricate and complex, are nevertheless encouraging."[92]

Another similar study investigating somewhat younger children between the ages of seven and ten found that output was higher for boys qualitatively, and higher for girls quantitatively.[93] McCall, Hogarty, and Hurlburt[94] found somewhat contradictory results. These researchers claim that girl infant test scores, in contrast to those of boys, showed higher correlation with later IQ. These findings suggest that, on the average, feedback makes for increasing convergence with age between IQ and creativity for boys but not for girls, although there are exceptions.

---

[92]Ibid., p. 440.

[93]R. M. Bhavnani and C. Hutt, "Divergent Thinking in Boys and Girls," Journal of Child Psychology and Psychiatry, 13 (1972), pp. 121-27.

[94]R. B. McCall, P. S. Hogarty, and N. Hurlburt, "Transition in Infant Sensorimotor Development and the Prediction of Childhood IQ," American Psychologist, 27 (1972), pp. 728-48.

In a careful re-analysis of a longitudinal study, Battle and Lacey[95] found that under the influence of maternal feedback, achievement was negatively correlated with active behavior in boys, and positively in girls. The researchers concluded that young mothers would respond with negative feedback to young boys who were overly active, but not to young girls. This, the researchers concluded, translated itself for boys into self-doubt, anxiety, and feelings of incompetence, as measured in later achievement scores. This finding seems to contradict popular lore that Western culture reinforces outgoing behavior in boys but not in girls, thus preparing boys better for adult competition and therefore achievement. In this study, too, "feedback" was equated with reinforcement and spanned a very large variety of overt behaviors.

There are a great many studies that focus on teacher feedback as the independent variable and sex-linked achievement output as the dependent variable for both boys and girls of varying ages. According to research, this cultural expectation translates itself into reinforcement by female teachers of conforming, unassertive behavior in boys, causing a concomitant drop in scholastic achievement. One study shows that teachers approve strongly of dependent, non-aggressive behavior in both boys and

---

[95] E. S. Battle and B. Lacey, "A Context for Hyperactivity in Children over Time," Child Development, 43 (1972), pp. 757-73.

girls,[96] making feedback contingent upon "approved" behaviors. Another study by Biber, Miller, and Dyer[97] found that while quantity of positive feedback is greater for girls, the quality of such feedback is equal for boys and girls. In terms of frequency of negative feedback, one study[98] verified that girls are seen more favorably by both male and female teachers, as measured by feedback in the form of spoken approval. In terms of perception, however, the boys saw themselves as being preferred by male teachers, while girls saw themselves as equally preferred by both. One might raise the question that while "spoken approval" might be relatively easy to be put to the empirical test, how would "perception" be measured? The two terms are not empirically equivalent, although the authors treat them as such.

It might be said that both age and sex typed behaviors seem to become more visible over time, and begin to be manifested in attitudes and behavior that aid or deter achievement. By fifth

---

[96] T. E. Levitin and J. D. Chanani, "Responses of Female Primary School Teachers to Sex Typed Behaviors in Male and Female Children," Child Development, 43 (1972), pp. 1309-16.

[97] H. Biber, L. B. Miller, and J. L. Dyer, "Feminization in Preschool," Developmental Psychology, 7 (1972), p. 86.

[98] P. C. Lee and A. L. Wolinsky, "Male Teachers of Young Children: A Preliminary Empirical Study," Young Children, 28 (1973), pp. 342-52.

grade, girls tend to underestimate their success, give up more easily than boys in the face of sustained failure,[99] and act according to what Dweck and Repucci[100] call "learned helplessness." These researchers found that internal responsibility for achievement is a function of academically successful boys and academically failing girls. Again, there are methodological difficulties here with an abstract such as "internal responsibility." The authors offer no adequate definition. The casual reader is tempted to equate this term with the more commonly used term "motivation."

The authors' findings were also seconded by Clifford and Cleary[101] who found that "internality" and academic performance correlated higher for boys than for girls, and the relationship between IQ and achievement was higher for girls than for boys. Levi and Tucker[102] present data in light of the above findings

---

[99] T. Gjesme, "Achievement-related Motives and School Performance for Girls," Journal of Personal Social Psychology, 26 (1973), pp. 131-36.

[100] C. S. Dweck and N. D. Repucci, "Learned Helplessness and Reinforcement Responsibility in Children," Journal of Personal Social Psychology, 25 (1973), pp. 109-16.

[101] M. M. Clifford and T. A. Cleary, "The Relationship between Children's Academic Performance and Achievement Accountability," Child Development, 43 (1972), pp. 647-55.

[102] P. Levy and J. Tucker, "Differential Effects of Streaming on Primary School Attainment," British Journal of Educational Psychology, 42 (1972), pp. 75-9.

suggesting that ability grouping (open feedback) might be particularly effective for boys, while disguised feedback in the form of heterogeneous grouping might benefit less capable boys more. House[103] presents evidence that boys welcome feedback in achievement-tied competition, while girls are more likely to avoid competition and hence are not as dependent on feedback. A related finding was that competition was sought out by boys with strong self-images, but not by girls with similar positive ego strength.

In summary, it is apparent that differential feedback makes for intricate and complex age, sex, and ability differences influencing achievement throughout the course of schooling. The availability, frequency, quantity, and quality of feedback appears to be the result of complex transactions between genetic and constitutional factors, cultural influences involving sex typing, and differential responses by significant adults, especially teachers. A clear-cut age, sex, and ability study pertaining to hierarchical learning under the influence of feedback is at this time not available.

---

[103]G. F. House, "Orientations to Achievement: Autonomous, Social Comparison and External" (PhD thesis, University of Michigan, Ann Arbor, 1972).

FEEDBACK AND HIERARCHICAL LEARNING

General Overview

Studies on hierarchical learning have been reviewed extensively by Holland[104] and Evans.[105] More recent but briefer reviews are given by Niedermeyer, Brown, and Sulzen[106] and Rayne, Krathwohl, and Gordon.[107] Tennyson[108] in his 1970 review incorporated an analytical approach into his summary. The most comprehensive and thoughtful review has been provided by Briggs,[109]

---

[104] J. G. Holland, "Research on Programming Variables," Teaching Machines and Programmed Learning. II: Data and Directions, ed. R. Glaser (Washington, D. C.: National Education Association, 1965), pp. 53-67.

[105] J. L. Evans, "Programming in Mathematics and Logic," Teaching Machines and Programmed Learning. II: Data and Directions, ed. R. Glaser (Washington, D. C.: National Education Association, 1965), pp. 68-90.

[106] F. Niedermeyer, J. Brown, and B. Sulzen, "Learning and Varying Sequences of Ninth-grade Mathematics Materials," Journal of Experimental Education, 37 (1969), pp. 61-66.

[107] D. A. Payne, D. R. Krathwohl, and J. Gordon, "The Effects of Sequence on Programmed Instruction," American Educational Research Journal, 4 (1967), pp. 125-32.

[108] R. D. Tennyson, "A Critical Review of Experimental Methodology in Instructional Sequencing" (Provo, Utah: Department of Instructional Research and Development, Brigham Young University, 1970), pp. 16-30.

[109] L. J. Briggs, "Sequencing of Instruction in Relation to Hierarchies of Competence," Monograph No. 3 (Pittsburgh, Pa.: American Institutes for Research, 1968).

particularly in view of hierarchical learning affecting sequential instruction.

Foremost advocates of hierarchical learning principles are Gagne,[110] Ausubel,[111] and Bruner.[112] Although their approaches and premises differ, the generic questions underlying their work might be phrased as follows: Is higher learning dependent upon the mastery of lower order tasks? If so, what might be done to augment effective hierarchical mastery? What conditions need to precede effective hierarchical learning? The study to follow was predicated upon the assumption that concise informational feedback is one such condition. An overview of the principles of hierarchical learning as espoused by these three theorists is therefore offered below.

Hierarchical learning, as espoused by Gagne, Ausubel, and Bruner, might be summed up as follows: Lower order tasks are believed to be components of higher order tasks, constituting a sequenced progression leading to increasingly complex performance. Pedagogically, one might condense the principles of hierarchical learning as follows: (1) No objective should be taught unless its

---

[110]Robert M. Gagne, "Learning Hierarchies," Educational Psychologist, 6 (1968), pp. 1-9.

[111]David P. Ausubel, Educational Psychology: A Cognitive View (New York: Holt, Rinehart and Winston, 1968).

[112]Jerome S. Bruner, Studies in Cognitive Growth (New York: Wiley, 1968).

prerequisite skills have been mastered. (2) There are a number of roads by which to master prerequisites. They can be taught by means of "linear instruction,"[113] in groups of "mathemagenical concepts,"[114] or by ". . . intuitively grasping and mastering sets."[115] Ausubel and Gagne might stress that tasks can be ordered in subordinate and superordinate relationships, hereby facilitating testing, and that individuals can and should be ordered with respect to their level of academic attainment. Bruner would disagree, at least in his earlier writings.

Gagne's Principles of Hierarchical Learning

Of the three theorists cited above, Gagne's work is probably best known as having given practical dimensions to theoretical structure. The concepts of learning hierarchies as

---

[113]Robert M. Gagne, The Conditions of Learning (New York: Holt, Rinehart and Winston, 1965).

[114]E. Rothkopf, "The Concept of Mathemagenic Activities," Review of Educational Research, 40 (1970), pp. 325-36.

[115]Jerome S. Bruner, "The Act of Discovery," Harvard Educational Review, 31 (1961), pp. 21-32.

developed by Gagne[116] and others[117] have recently acquired new prominance in instructional design. Gagne's complete treatise regarding categorization of tasks according to learning requirements is offered in his book, The Conditions of Learning.[118] Here, Gagne describes eight varieties of learning, ranging from simple operant conditioning to very complex problem solving tasks. Categories described are: conditioned responses; chaining; verbal associations; sets of discriminations; concept formation; rule learning; and the ability to use discriminations, rules, and concepts to solve problems.

Gagne differentiates between the learning of verbal information, intellectual skills, and cognitive strategies, but he believes that the content and sequence of a curriculum can be based upon a rational analysis,[119] as does Gibson[120] who classified the

---

[116] Robert M. Gagne, "Learning Hierarchies," Educational Psychologist, 6 (1968), pp. 1-9.

[117] L. B. Resnick, M. C. Wang, and J. Kaplan, "Behavior Analysis in Curriculum Design: A Hierarchically Sequenced Introductory Mathematics Curriculum," Monograph 2, ED 047 954 (Pittsburgh: Learning Research and Development Center, 1970).

[118] Robert M. Gagne, The Conditions of Learning (2nd ed.; New York: Holt, Rinehart and Winston, 1970).

[119] Robert M. Gagne, "Curriculum Research and the Promotion of Learning," Perspectives of Curriculum Evaluation (New York: Rand McNally, 1967), p. 74.

[120] Eleanor J. Gibson, "Learning to Read," Science, 148 (1965), pp. 1066-72.

behavioral meaning of the hierarchically ordered stages. A very complex analytical attempt at breaking down a problem situation into its stimulus component parts has been attempted by Gagne. He divides learning stimuli into three groups: (1) integral stimuli, (2) activating stimuli, and (3) stimuli from recall. He further subdivides these groups as follows: Integral stimuli fall into (a) task stimuli and (b) context stimuli. Activating stimuli have four subdivisions--those that (a) give directions, (b) activate recall, (c) activate a set, and (d) arouse a motivational state. Stimuli from recall fall into (a) information recall, (b) intellectual skills recall, (c) reinstatement of a performance set, and (d) cognitive strategies. From a summary article by Gagne, these hard-to-grasp divisions and sub-divisions might be paraphrased more clearly and comprehensively as follows:

> Feedback, by giving orienting directions, activates a predisposition called a set which initiates a selective search strategy that determines how the learner will organize previous knowledge to problems yet untried.[121]

Many studies by Gagne and others go beyond simple recall and address themselves to "rule learning" or the application of principles to problems following the learning. Representative examples are an early mathematical study by Gagne and Brown[122] and

---

[121] Gagne, op. cit., p. 15.

[122] R. M. Gagne and L. T. Brown, "Some Factors in the Programming of Conceptual Learning," Journal of Experimental Psychology, 62 (1961), pp. 313-21.

Brown.[123] A partial finding was that hierarchical learning may have little effect on retention of simple verbal learning, but demonstrable effect on learning cognitive strategies. Merrill, Barton, and Wood,[124] in another experiment designed to promote mastery, gave a step-by-step solution to each missed problem, contrasted with simple right-wrong feedback. Informational feedback proved superior to the more vague and disguised "right-wrong" approach. One gets the feeling with this study, as with the preceding three, that recent years have seen an increasing attempt to define feedback operationally rather than impressionistically.

Ausubel's Principles of Hierarchical Learning

Ausubel,[125] much like Gagne, emphasizes the importance of meaningful context to learning and retention. Ausubel is best known

---

[123] J. L. Brown, "Effects of Logical and Scrambled Sequences in Mathematical Materials on Learning with Programmed Instruction Materials," Journal of Educational Psychology, 61 (1970), pp. 41-45.

[124] M. D. Merrill, K. Barton, and L. E. Wood, "Specific Review in Learning a Hierarchical Imaginary Science," Journal of Educational Psychology, 61 (1970), pp. 102-109.

[125] David P. Ausubel, "The Use of Advance Organizers in the Learning and Retention of Meaningful Verbal Material," Journal of Educational Psychology, 51 (1960), pp. 267-72.

for his work on receptive hierarchical verbal learning,[126] the characteristics of which are sketched below:

There are organizational principles by which human beings acquire and retain stable knowledge. These principles can be discovered only through an applied engineering type of research. "Ideational scaffolding" (more commonly known as advance organizers) seems to be one such organizational principle. The learner employs a "set" in order to incorporate into his cognitive structure potentially meaningful materials which are "subsumed" by established entities within that structure. Sequential and/or hierarchical learning constitutes relevant ideational scaffolding. Each new increment of learning serves as an anchoring point for subsequent learning. It is very important, according to Ausubel, to strengthen lower order learning by adequate feedback in order to facilitate the mastery of higher order tasks.[127]

Recall of information, for example, presupposes the need for a considerable amount of meaningfully organized data previously learned. There has to have been prior learning of subordinate rules

---

[126] David P. Ausubel, The Psychology of Meaningful Verbal Learning (New York: Grune and Stratton, 1963).

[127] David P. Ausubel, "The Use of Advanced Organizers in the Learning and Retention of Meaningful Verbal Materials," Journal of Educational Psychology, 51 (1960), pp. 267-72.

and concepts much in the manner described by Gagne.[128] A "meaningful learning set," according to Ausubel, is much more than a simple anchoring device. It strengthens a disposition to learn or perform in a particular way. It reflects the influence of recent past learning on subsequent learning. According to Ausubel, it is both the manifestation of a methodological skill and the manifestation of an appropriate performance attitude.[129] He calls this readiness the "warm-up effect."[130] Both methodological skill and performance attitude contribute to positive hierarchical transfer.

Ausubel holds feedback to be an extremely important though not indispensable variable in hierarchical learning. To quote:

> ...Behaviorally oriented theorists tend to attribute the effects of feedback solely to reinforcement, thus equating knowledge of results with other kinds of rewards for learning...Awareness that the results of learning will be made available constitutes an incentive condition, thereby enhancing the strength of the underlying devices. But the facilitating effects of feedback are hardly exhausted by these reinforcement and motivational mechanisms. Knowledge of results has other purely cognitive effects on learning. It confirms appropriate meanings and associations, corrects errors, clarifies misconceptions, and indicates the relative adequacy

---

128 Robert M. Gagne, The Conditions of Learning (2nd ed.; New York: Holt, Rinehart and Winston, 1970).

129 David P. Ausubel, "Cognitive Theory of School Learning," Psychology in Schools, 6 (October, 1969), pp. 331-5.

130 David P. Ausubel, Educational Psychology: A Cognitive View (New York: Holt, Rinehart and Winston, 1968), p. 53.

with which different portions of the learning task have been mastered. Thus, as a result of the feedback he receives, the subject's confidence in the validity of his learning products is increased...[131]

The usefulness of Ausubel's "performance sets" has also been studied by other researchers. Frase[132] speaks of "orienting directions," while Rothkopf[133] has coined the phrase, "mathemagenic effects." Frase investigated feedback in the form of directed questions. In this study, questions were seen as "activating a set to confirm and thereby reinforce valuable mathemagenic behaviors," having the effect of causing more generalized learning and retention.[134] A similar study[135] showed that providing the learner with information about the organized structure of passages aided recall. Frase actually speaks of a "cybernetics effect," using a vocabulary familiar to a management principle used in administration.

---

[131] Ibid., p. 118.

[132] L. T. Frase, "Boundary Conditions for Mathemagenic Behaviors," Review of Educational Research, 40 (1970), pp. 337-47.

[133] E. Z. Rothkopf, "The Concept of Mathemagenic Activities," Review of Educational Research, 40 (1970), pp. 325-36.

[134] L. T. Frase, "Paragraph Organization of Written Materials: The Influence of Conceptual Clustering upon the Level and Organization of Recall," Journal of Educational Psychology, 60 (1969), pp. 394-401.

[135] L. T. Frase, "Cybernetic Control of Memory While Reading Connected Discourse," Journal of Educational Psychology, 60 (1969), pp. 49-55.

Bruner's Principles of Hierarchical Learning

Bruner's[136] approach to hierarchical learning can be described as wholistic and conceptual. Bruner stresses that if a learner grasps the structure of a subject, he understands it in a way that permits many other learning tasks to be related to it meaningfully. Bruner feels that in conceptual learning, there are sets that consist of rules, principles, hypotheses, laws, and even models and paradigms. If these sets are grasped early, they will be self-reinforcing to the learner, and progressive higher-order learning will result.[137] Additional data regarding wholistic learning were also provided by Eisenberg and Walbesser[138] and by Gagne.[139] These studies are specific to mathematics transfer.

According to Bruner, learning cannot be translated into a generic form until there has been enough mastery of the specifics of a situation to permit the discovery of lower-order generalities, which

---

[136] Jerome Bruner, The Process of Education (Boston: Harvard University Press, 1962).

[137] Jerome Bruner, "The Act of Discovery," Harvard Educational Review, 31 (1961), pp. 21-32.

[138] T. A. Eisenberg and H. H. Walbesser, "Learning Hierarchies--Numerical Considerations," Journal for Research in Mathematics Education, 2 (1971), pp. 244-56.

[139] Robert M. Gagne, "Factors in Acquiring Knowledge of a Mathematical Task," Psychological Monographs, 76 (1962), No. 526.

can then be re-combined into more generic coding systems. Bruner however, stresses discovery of regularities rather than the teaching and reinforcing of these regularities. He would prefer inductive teaching, while Gagne and Ausubel would take a more systematic, deductive approach.

Related Research Pertaining to Hierarchical Learning

These three basic approaches by Gagne, Ausubel, and Bruner have stimulated a flurry of studies on hierarchical strategies. Homme[140] showed that the principles of contingency and feedback had broad applicability in classroom activities. Wolf and Risley[141] stated that desired behaviors could be increased by sequencing feedback hierarchically. In studies on "prompting"--a form of weaker, "disguised feedback"--Holland[142] and Sidman and Stoddard[143] reviewed instructional programs and commented on the desirability

---

[140] L. Homme, How to Use Contingency Contracting in the Classroom (Champaign, Ill.: Research Press, 1969).

[141] M. M. Wolf and T. R. Risley, "Reinforcement: Applied Research," The Nature of Reinforcement, ed. R. Glaser (New York: Academic Press, 1971), pp. 310-25.

[142] J. G. Holland, "Research on Programming Variables," Teaching Machines and Programmed Learning. II: Data and Directions (Washington, D.C.: National Education Association, 1965).

[143] M. Sidman and L. T. Stoddard, "The Effectiveness of Fading in Programming a Simultaneous Form Discrimination for Retarded Children," Journal of Experimental Analysis of Behavior, 10 (1967), pp. 3-15.

and applicability of prompting. Hierarchical learning was found to be augmented by "context feedback."[144] Bruner, Goodnow, and Austin[145] talk about "thinking strategies" that presuppose knowledge of prior results. Gagne[146] distinguishes hierarchical variables as affecting verbal information, intellectual skills, and cognitive strategies differently. Rothkopf[147] has investigated the establishment of performance sets in a temporal setting. Bem[148] found that once subordinate skills were learned, previously unpracticed higher-order skills would be mastered with relative ease.

These studies, though often specific and somewhat narrow in focus, seem to have overcome many of the earlier methodological difficulties such as poor refinement of the independent and dependent variables, inadequacy of measuring devices, and poor sample size.

---

[144] W. D. Rohwer, Jr., and J. R. Levin, "Action, Meaning, and Stimulus Selection in Paired Associate Learning," Journal of Verbal Learning and Verbal Behavior, 7 (1968), pp. 137-41.

[145] J. S. Bruner, J. J. Goodnow, and G. A. Austin, A Study of Thinking (New York: Wiley, 1956), p. 49.

[146] R. M. Gagne, "Domains of Learning," Interchange, 3 (1972), pp. 1-8.

[147] E. Z. Rothkopf, "The Concept of Mathemagenic Activities," Review of Educational Research, 40 (1970), pp. 325-36.

[148] S. L. Bem, "Verbal Self-control: The Establishment of Effective Self-instruction," Journal of Experimental Psychology, 74 (1967), pp. 485-91.

The influence of operational principles is felt, and the target areas are for the most part in the instructional, curricular realm.

Some writers have also concerned themselves with the validation of learning hierarchies.[149] Gagne and Rohwer[150] reviewed a number of transfer studies. Shulman[151] gave an analytical and comparison view and emphasized contrasting issues brought out by the work of Ausubel, Bruner, and Gagne. Bruner[152] in turn brought his theoretical views on hierarchies to bear upon learning practices, as did Wohlwill,[153] Merrill,[154] and Tennyson and

---

[149] L. B. Resnick and M. Wang, Approaches to the Validation of Learning Hierarchies, Preprint 50 (Pittsburgh: Learning Research and Development Center, University of Pittsburgh, 1969).

[150] R. M. Gagne and W. D. Rohwer, Jr., "Instructional Psychology," Annual Review of Psychology, 20 (1969), pp. 381-418.

[151] L. S. Shulman, "Psychology and Mathematics Education," Mathematics Education, ed. E. G. Begle. Sixty-ninth Yearbook of the National Society for the Study of Education (Chicago: NSSE, 1970).

[152] J. S. Bruner, "The Skill of Relevance and the Relevance of Skills," Saturday Review, April 18, 1970, pp. 66-8, 78-9.

[153] J. F. Wohlwill, "The Place of Structured Experience in Early Cognitive Development," Interchange, 1 (1970), pp. 13-27.

[154] M. D. Merrill, "Necessary Psychological Conditions for Defining Instructional Outcomes," Educational Technology, 2 (1971), pp. 34-9.

Merrill.[155] No fully satisfying method of hierarchy validation has yet come to light.

SUMMARY

One might conclude that there are a number of provocative claims pertaining to feedback per se and with regard to ability, age, and sex, as well as feedback with regard to hierarchical learning. There is, however, as yet no consensus of what actually constitutes hierarchical mastery or what might be the means to reach or teach mastery. One pattern, however, seems clear: Once a learner has acquired a response to a particular stimulus situation, a similar but more difficult situation will more easily elicit the behavior that has been previously learned. In other words, once lower order behaviors have been reinforced adequately, the probability of higher order behavior in a similar but more difficult context will be increased. This attribute of behavior is called generalization. In part, the study to follow was predicated upon the assumption that informational feedback to lower order learning would carry generalized behaviors into higher order tasks, showing results in immediate achievement.

---

[155] R. D. Tennyson and M. D. Merrill, "Hierarchical Models in the Development of a Theory of Instruction: A Comparison of Bloom, Gagne, and Merrill," Educational Technology, 2 (1971), pp. 27-31.

Chapter 3

METHODS OF THE STUDY

INTRODUCTION

In this chapter, the scope and sequence of the research investigation are described. A rationale for choosing the WISC-R as a learning experiment with built-in measurement properties is given. The research design is outlined. A comprehensive description of the WISC-R is offered, its statistical properties are listed, and a precise description of the standard test procedures (administered to the control group) and feedback modifications (administered to the experimental group) are given. The reasons and methods of identifying the sample are specified. The details of the data collection are outlined. Finally, the statistical tools are specified.

In order to recoup the essence of Chapter 1 and to focus the study, the research problem statements are repeated:

First, what is the evidence regarding learning as a function of informational feedback in global, Gestalt-type task mastery where feedback, given in response to lower order learning, may aid the mastery of higher order tasks?

Second, what evidence exists that, given informational feedback, students will suffer depletion of motivation as a result, as inferred by a decreased quality and quantity of learning?

Third, do entry variables such as ability, age, and sex interact with feedback in a significant way?

For additional clarity, the hypotheses are also repeated:

Ho_1: There is no difference between the experimental and control groups on the following WISC-R scales: Information, Similarities, Arithmetic, Vocabulary, Comprehension, Picture Completion, Picture Arrangement, Block Design, Object Assembly, and Coding.

Ho_2: There is no difference between the experimental and control groups on the verbal and performance summation scales.

Ho_3: There is no difference between the experimental and control groups on the full scale score.

Ho_4: The effects of feedback are constant with respect to ability.

Ho_5: The effects of feedback are constant with respect to age.

Ho_6: The effects of feedback are constant with respect to sex.

The problems and hypotheses stated presuppose (1) hierarchical learning tasks subject to informational feedback,

(2) precise quantification of "learning" in order to assess whether or not feedback aids or deters learning, and (3) strict controls on the entry variables of ability, age, and sex. The WISC-R, individually administered to 30 sets of identical twins, provided all three experimental prerequisites as outlined above.

THE WISC-R AS A TEST AND/OR LEARNING EXPERIMENT

The purpose of this experiment was to apply the WISC-R as a series of "novel" hierarchical learning tasks with built-in measurement properties. The three WISC-R summation scales as well as the ten individual subscales can be translated easily and quickly into norm-referenced, quantified data. The experiment was set up in such a way that the control group scores would yield IQ's, but only "incidentally." The experimental scores, containing a feedback component, could be compared statistically to the data from the controls.

The WISC-R scales were chosen, in part, for their proven appeal due to their organized diversity.[1] The test offered assured uniformity of both the initial learning stimuli (test items) as

---

[1]Oscar Krisen Buros (ed.), The Seventh Mental Measurement Yearbook (Highland Park: The Gryphon Press, 1972), p. 800.

well as uniformity of treatment (feedback in the form of answers to the test items). The pre-tested hierarchical order of difficulty of the questions and tasks, such as only a standardized test can provide, yielded high quality raw data which could be translated to finely divided age norms derived from excellent statistical bases. Contrary to most experimental studies, where the treatment is carefully isolated but the actual learning task is poorly controlled, in the present study it was possible to have tight control of both treatment and task.

RESEARCH DESIGN

This study used an experimental group-control group matched subjects design, with one modification: treatment and testing occurred simultaneously. Despite its apparent simplicity, this study relied upon the triple strength of (1) a relatively large twin sample with built-in controls as to ability, age, and sex, (2) a strictly controlled and circumscribed behavioral setting, and (3) the advantage of translation of behavioral raw scores into norm-referenced quantification.

As outlined more fully in the individual subscales description given in Appendix A, the control group took the WISC-R according to standardized instructions. The experimental group took the identical items, followed up by feedback as to the correctness or incorrectness of their responses. At the outset, the experimental subjects were told: "This test is made for children ranging from kindergarten all the way through high school. There is no way you can do it all. Do as well as you can, but if you cannot solve an item, I am going to give you the answer."

Feedback as the independent variable was circumscribed as follows: Any correct response was followed by "Good!" or "Right!" If an item was failed, or was only partially credited, the subject was given a "good" answer as provided by the manual, and as detailed in Appendix A.

The dependent variables consisted of the scores of the WISC-R. Hierarchical learning as a function of concise informational feedback was investigated by a comparison of the ten individual subscales of the experimental group and control group. Whether or not students suffered a decreased quantity or quality of learning as a result of informational feedback was tested by a comparison of the three summation scales. The interactional effects of feedback and ability, age, and sex were tested by comparing the subscales and summation scales of young versus older children, bright versus slow children, and boys versus girls.

INSTRUMENTATION

General Description of the WISC-R

According to the author David Wechsler,[2] the WISC-R is designed and organized as a test of general intelligence. It was standardized on 2200 subjects, using stratification measures to reflect ethnic as well as demographic population proportions found in the 1970 United States Census. It is used primarily as a diagnostic tool for predictive purposes, but it has also found application in clinical use.[3]

The WISC-R is a norm-referenced test; the scores obtained are deviation scores. A full scale score of 100 is set equal to the mean total score for each age, and the standard deviation is set equal to 15 points. The twelve subscales that compose this test carry a mean of 10 and a standard deviation of 3. Only ten of these subscores are commonly used. They fall into five verbal-intellectual and five spatial-performance tests, as verified by numerous factorial

---

[2] David Wechsler, Manual for the Wechsler Intelligence Scale for Children - Revised (New York: The Psychological Corporation, 1974).

[3] Alan J. Glasser and Irla L. Zimmerman, Clinical Interpretation of the Wechsler Intelligence Scale for Children (New York: Grune and Stratton, 1967).

studies.[4] This dichotomy is used as a way of identifying two principal modes by which human abilities seem to express themselves. Any test administration will yield thirteen scores. Ten scores tap various cognitive sub-areas. The three summation scores are commonly translated into IQ's.

The range of the test is from 6 years 0 months through 16 years 11 months. Raw scores for all ages can be converted to standardized age norms easily and rapidly. Norms for both girls and boys are the same.

The items are interesting to children and hold their attention well. The scales are easily yet strictly administered, assuring great uniformity of presentation. The Manual and protocols give clear directions, are of convenient size, easy to handle, and easy to score. Average testing time ranges from 50 to 75 minutes.

Standard Test Procedures

The WISC-R can only be administered and scored by a trained examiner.[5] The tester is requested to follow all Manual directions meticulously, and to record and evaluate student responses

---

[4]A. E. Maxwell, "A Factor Analysis of the Wechsler Intelligence Scale for Children," British Journal of Educational Psychology, 29 (1959), pp. 237-41.

[5]The investigator had much practice with the WISC-R, having administered it more than a thousand times during six years of employment as a school psychologist.

quickly and expeditiously. The intention of the battery is to assess the child's performance under a fixed set of conditions, <u>not</u> to test the limits of his knowledge. All performance tests and one verbal test have time limits specifying seconds. A precision stop watch is used inconspicuously.

It is important for the examiner to have mastered, if not memorized, the details of administration and scoring, permitting maximum attention to the child. The examiner is encouraged to reward effort, though not necessarily success. No dissatisfaction with any response the child has given may be shown, nor may the child expect approval for correct responses. In other words, during standard testing procedures the child remains largely "in the dark" regarding the appropriateness of the answer he has supplied.

The test sequence is fixed, with verbal/non-verbal subtests alternating. Starting item points are given for each age group. The test is discontinued after a certain number of consecutive failures.

Strict directions for scoring test items are provided in the Manual. The rules for most WISC-R items are objective and do not require subjective "guesswork." If for some reason one of the subtests has been spoiled, or has not been given or completed, it is possible to pro-rate the rest of the obtained data.

Description of the Subscales

1) Thirty items make up the subscale Information. This test is in part an index of good memory, cultural environment, alertness, wide reading, and ambition. A primary function covered is long-range retention of "everyday facts." Representative questions and answers are:

Q. "Who discovered America?" A. "Christopher Columbus."

Q. "Which month has one extra day during leap year?"

A. "February."

Q. "Name the two countries that border the United States."

A. "Canada and Mexico."

Scoring is done on a pass/fail basis. Testing is discontinued after five consecutive failures.

2) Sixteen items make up the subscale Similarities. This test is in part an index of concept formation and the ability to generalize from experience. Concepts may be concrete, abstract, perceptual, or functional. A primary function is to detect similarities in dissimilar items or concepts. Representative test items are:

Q. "How are an apple and a banana alike?" A. "They are both fruit."

Q. "How are a telephone and a radio alike?" A. "They are means of communication."

Q. "How are liberty and justice alike?" A. "They are social ideals."

Scoring is done as follows:

Full credit = 2 points. Partial credit = 1 point. No credit = 0 points. Testing is discontinued after three consecutive failures.

3) Eighteen items make up the subscale Arithmetic. This test is in part an index of concentration, attention, the ability to perform mental operations, and the ability to work mentally without concrete aids. The primary function covered is computational skills. Representative test items are:

Q. "If I cut an apple in half, how many pieces will I have?" A. "Two."

Q. "If you buy two dozen pencils at 45 cents a dozen, how much change should you get back from one dollar? A. "10 cents."

Q. "If three pieces of bubble gum cost 5 cents, what will be the cost of 24 pieces?" A. "40 cents."

Scoring is done on a pass/fail basis. Testing is discontinued after three consecutive failures.

4) Thirty-two items make up the subscale Vocabulary. This test is in part an index of verbal intelligence and learning ability, familial cultural background, abstract thinking ability, and long-range richness of ideas. The primary function covered is expressive rather than receptive language. Representative test items are:

Q. "What is a clock?" A. "It tells what time it is."

Q. "What do we mean by prevent?" A. "To keep something from happening."

Q. "What is an affliction?" A. "A burden you must bear."

Scoring is done as follows: Full credit = 2 points. Partial credit = 1 point. No credit = 0 points. Testing is discontinued after five consecutive failures.

5) Seventeen items make up the subscale Comprehension. This test is in part an index of social maturity, social judgment, practical intelligence, ability to verbalize well, knowledge, and practicality. The primary function covered is understanding reasons for everyday experience. Representative test items are:

Q. "What is the thing to do if you lose a ball that belongs to one of your friends?" A. "You buy him a new one."

Q. "Why do we have to put stamps on letters?" A. "To pay for the mail delivery."

Q. "Why is it good to hold elections by secret ballot?" A. "To protect your privacy in voting and prevent you from being pressured to vote a certain way."

Scoring is done as follows: Full credit = 2 points. Partial credit = 1 point. No credit = 0 points. Testing is discontinued after four consecutive failures.

6) Twenty-six items make up the subscale Picture Completion. This test is in part an index of visual memory and visual alertness, figure-ground perception, and good concentration. Primary functions covered are the ability to establish a learning set and to rule out background stimuli in order to focus on detail.

The subjects are asked to react to cards displayed for 20 seconds each, and are asked to tell "... what is missing." Missing items may be the whiskers on a cat, an eyebrow in a profile, or the shadow of a tree, for example.

Scoring is done on a pass/fail basis. Testing is discontinued after 4 consecutive failures.

7) Twelve items make up the subscale Picture Arrangement. This test is in part an index of alertness to detail, forethought, planning ability, high sequential thought processes, and social skill. Its primary function is to perceive cause-and-effect relationships very quickly and to note detail pertaining to time.

The subjects are given sets of cards in disarray order and are asked to "... put them in order." The experimental subjects, after failing an item, were told:

"You missed this one. Watch me. The order should have been like this."

The correct order was then demonstrated, and the subjects were permitted to study the sequence for fifteen seconds.

Scoring was done according to the instructions listed on the protocol sheet. The individual items were second-timed, with special "bonuses" given for speed. The amount of credit, therefore, differed from item to item.

8) Eleven items make up the subscale <u>Block Design.</u> This test is in part an index of conceptualization, analytic synthesizing, trial and error flexibility, and non-verbal reasoning. Two functions of the test are three-dimensional space perception and eye-hand coordination.

The subjects were given a set of blocks and were asked to reproduce designs utilizing red and white patterns. The experimental subjects, upon failing an item, were told:

"You missed this one. Watch me. It should have looked like this."

The correct pattern was demonstrated, and the subjects were permitted to study the pattern for 15 seconds.

Scoring was done according to the instructions on the protocol sheet. The individual items were second-timed, with special bonuses given for speed. The amount of credit, therefore, differed from item to item.

9) Four items make up the subscale <u>Object Assembly.</u> The test is in part an index of trial and error adaptability, flexibility to explore new solutions quickly and expeditiously, problem solving

techniques, and the ability to organize an approach around an idea or plan. Two functions covered are eye-hand coordination and pattern recognition.

The subjects were asked to close their eyes while puzzle pieces were arranged in a pre-determined, specified manner. They were then asked to put the pieces together to make a whole out of parts.

The experimental subjects, upon failing an item, were told:

"You missed this one. Watch me. It should have looked like this."

The correct assembly was demonstrated, and the subjects were permitted to study the pattern for 15 seconds.

Scoring was done according to the instructions on the protocol sheet. The individual items were second-timed, with special "bonuses" given for speed, and with partial credit given for partial completion.

10) The subscale Coding does not have individual items. It is in part an index of high motivation and dexterity, of memorization of symbols, and of the ability to work well under pressure. The primary function is to substitute one symbol system for another correctly and speedily.

No feedback is possible for this score, due to the nature of the task. It has special strategic emphasis in that it comes at the end of a relatively strenuous, anxiety-producing test. It is therefore given additional clinical significance above and beyond the rest of the subscales.

Scoring is done as follows: The correct symbols are counted and the final raw score is the number of symbols so reproduced within 120 seconds.

Statistical Test Properties of the WISC-R

The reliability coefficients of the individual tests, and of the verbal, performance, and full scale scores, are presented in Appendix B. The average full scale WISC-R reliability coefficient is .96. For the verbal scale it is .94, and for the performance scale it is .90.

Standard error of measurement scores for the WISC-R are presented in Appendix C for all age groups. The standard error of measurement indicates the band of error surrounding a test score. The average SE_m for the full scale score is 3.19. For the verbal scale it is 3.60, and for the performance scale it is 4.66.

Wechsler IQ's are fairly close numerically to IQ's of other well-standardized tests. Correlation studies show the relationship between the WISC-R and other individually administered tests to be as follows:

WPPSI	-	WISC-R	.82
WAIS	-	WISC-R	.95
Stanford-Binet	-	WISC-R	.73

POPULATION AND SAMPLE

The subjects in this study consisted of 30 sets of twins ranging in age from 6 to 16. The primary criterion for selection was that they be identical, assuring relative genetic constancy. Findings pertaining to genetic constancy in twins are cited below in support of the rationale for choosing twins non-randomly since it was felt that twinning would strengthen the conclusions as to the magnitude of learning measured in this study.

Monozygotic twinning occurs randomly in three out of every 1,000 births, and follows no discernible genetic patterns.[6] Identical twins are alike in sex and age and are believed to be near-alike in cognitive capacity, as verified by numerous correlational studies. These studies have investigated ability as measured by either conventional intelligence tests or developmental tests. Wilson and Harpering[7] found ability correlations to be between .81 and .85 for identical twins and .54 to .74 for fraternal twins.

---

[6] Encyclopedia Britannica, Volume 15 (Chicago, 1971), p. 985.

[7] R. S. Wilson and E. B. Harpering, "Mental and Motor Development in Infant Twins," Developmental Psychology, 7 (1972), p. 277.

A repeat of this study by Wilson[8] with a larger sample and a lowering of ages to three, six, nine, and twelve months yielded similar results, with correlations fluctuating from .81 to .84 for identical twins, and from .61 to .74 for fraternal twins. A study by Erlenmeyer-Kimling and Jarvik[9] of children of school age cites median correlations for fraternal twins as .53, for identical twins reared together as .87, and for identical twins reared apart as .75. Burt[10] reported correlations of .88.

Jensen summarized four of the more important studies, reporting a combined value of .85 to be representative and pointing out that an *r* of .37 would correspond to a 6 point variation on an IQ scale such as the WISC-R.[11] In addition, Jensen cites correlations for intelligence and scholastic assessments as follows:

---

[8]R. S. Wilson, "Twins: Mental Development in the Preschool Years," Developmental Psychology, 10 (1974), p. 580.

[9]L. Erlenmeyer-Kimling and L. F. Jarvik, "Genetics and Intelligence: A Review," Science, 142 (1963), p. 1477.

[10]C. Burt, "The Genetic Determination of Monozygotic Twins Reared Together and Apart," British Journal of Psychology, 57 (1966), pp. 137-53.

[11]A. R. Jensen, Genetics and Education (New York: Harper and Row, 1972), p. 313.

	Identical Twins Reared Together, N = 83	Identical Twins Reared Apart, N = 30
Intelligence		
Group Test	.944	.771
Individual Test	.921	.843
Final Assessment	.925	.876
Mean	.929	.829
Scholastic		
General	.898	.681
Reading and Spelling	.944	.681
Arithmetic	.862	.723
Mean	.900	.683

The strengths of the above statistics were seen as justification for making the criterion of "identical twins" paramount in subject selection.

Given the relative rarity of identical twinning, no attempt was made to match demographic factors pertaining to the subjects' parents. All children lived in or within the surroundings of Stockton, California, a city with a population of approximately 120,000. The social climate may be described as that of a small suburban California town.

The breakdown of recruitment was as follows:

Lincoln Unified School District	18
Stanislaus Union School District	2
Mothers of Twins Club	7
Self-referrals	3

The breakdown according to ability was as follows:

Table 1

Breakdown of Sample by Control Subjects' IQ

IQ	N	
83-89	4	
90-94	4	
95-99	4	$\overline{X}$ ("slow") = 94
100-104	2	Range: 83-105
105-109	3	
110-114	2	$\overline{X}$ ("bright") = 112
115-119	4	Range: 106-133
120-124	2	
125-133	5	

The breakdown according to age was as follows:

Table 2

Breakdown of Sample by Age

Age	N	
6	2	
7	6	
8	6	
9	4	
10	3	$\overline{X}$ ("young") = 7.4 years
11	2	
12	2	
13	2	$\overline{X}$ ("old") = 11.6 years
14	1	
15	1	
16	1	

Note: Tables 1 and 2 show arbitrary organismic variable divisions on ability and age by ranking the subjects and using the median as the point of division.

The breakdown according to sex was as follows:

Males	17 pairs
Females	13 pairs

DATA COLLECTION

The twin sets drawn from the Lincoln Unified School District were recruited as follows: The researcher initially contacted the respective central administrative personnel and asked for a list of names and ages of identical twins in the district. A letter was then sent to the parents from the central office indicating the nature of the study and stating that their children's participation was sought, but not mandatory, in legitimate academic research. Parents returned a card indicating by "yes" or "no" whether they wished their children to be included. Of 21 parents so contacted, 18 chose to participate.

The researcher then contacted these parents by telephone, and a time and place for testing was arranged. If parents requested it, a home visit was made prior to testing, during which the purpose of the study was explained. Parents were told that the siblings would not "compete" with each other--that different teaching strategies rather than comparisons were being explored. This proved to be very reassuring to parents, many of whom seemed to fear an ability comparison. A follow-up visit was also made on request, and the obtained data were shared with the parents. This care with the "public relations" aspect paid off in very good participation and very high cooperation on the part of both parents and school district personnel.

The sets of twins from the "Mothers of Twins Club" were drafted during one of their regular monthly meetings. All parents of identical twins present chose to participate. Three mothers whose twins attended the Stockton Unified School District had heard of the study and asked to have their children included. Two sets of twins came from a neighboring school district (Stanislaus Union School District) and were chosen to balance out a gap in age distribution.

All testing was done out of school time, mostly in late afternoons and on week-ends. With younger children it was sometimes possible to test in early morning. All testing was done at the experimenter's office, with two exceptions, where testing was done in the parents' home. Parents were never present during the experiment. All data were collected during the period of October 1978 and March 1979.

Prior to the start of the experiment, the WISC-R was administered to five sets of fraternal twins as a "pilot" in order to attempt to pinpoint any cursory difficulty areas in test administration, and in order to assess whether there was a large time discrepancy between the control group taking the WISC-R "straight" and the experimental group who were given informational feedback. The average time for the controls was 71 minutes, and for the experimentals 80 minutes. No observable difficulties in test-taking or treatment were noted.

STATISTICAL ANALYSIS

The two-tailed *t*-test (or an approximation to the *t*-test where unequal variances were assumed) was chosen as the appropriate statistical tool for all data, with the exception of a two-way and three-way interaction analysis between feedback and ability, age, and sex. An analysis of variance was performed on these data, requiring a slightly reduced sample (N = 28) to assure cell equivalency, a desirable prerequisite for this particular statistical treatment.

The results of the experiment are presented in the tables in Chapter 4. Interpretation of the results, and conclusions and recommendations drawn from the data, are organized and summarized in Chapter 5.

SUMMARY

Chapter 3 described the scope and sequence of the research investigation. It also cited statistical data in support of an experiment combining the relative genetic constancy inherent in identical twins with the relative experimental constancy assured by a standardized test such as the WISC-R. Finally, methods of sampling, data collection, and statistical treatment were specified.

Chapter 4

FINDINGS OF THE STUDY

INTRODUCTION

Three interrelated goals of this study were to find empirical answers to three questions: 1). Does feedback, given in response to lower order learning, facilitate the mastery of higher order tasks? 2). Will students under the influence of feedback suffer depletion of motivation as a result, as inferred by a decreased quality and quantity of learning? 3). Do student entry variables such as ability, age, and sex interact with feedback in a significant way? Chapter 4 offers statistical evidence for the acceptance or rejection of six null hypotheses extrapolated from these questions. The tables are ordered as follows: Table 3 is a statistical summary of Hypothesis One. Table 4 answers Hypothesis Two. Table 5 is in response to Hypothesis Three. Hypotheses Four, Five, and Six are broken into two parts: Tables 6, 7, and 8 pertain to questions on main effects between the interaction of feedback and ability, age, and sex, and Tables 9 and 10 present the significant two-way interaction effects between feedback and student characteristics.

HYPOTHESIS ONE

Ho_1 reads as follows: There is no difference between the experimental and control groups on the following WISC-R scales: Information, Similarities, Arithmetic, Vocabulary, Comprehension, Picture Completion, Picture Arrangement, Block Design, Object Assembly, and Coding. This hypothesis addresses itself to the nature of hierarchical learning as evidenced by the partition of the scales, as well as to the question of directionality of gain or loss. To paraphrase a point made earlier: The hierarchical WISC-R scales are assumed to tap relatively discrete modes of cognitive functioning. The degree to which feedback influenced these different learning tasks, as well as the direction of this influence, are answered in Table 3 in response to Hypothesis One.

Table 3

Comparison between the Experimental and Control Groups' Scaled Scores on Ten WISC-R Subscales (N = 30 for each group)

Scale	Group	Mean	Standard Deviation	Standard Error	(Difference) Mean	t-Value	Degrees of Freedom	2-Tail Prob.
Information	Exp.	10.6333	2.671	0.488	-0.6667	-2.16	29	0.039 *
	Cont.	11.3000	3.365	0.614				
Similarities	Exp.	12.4000	2.884	0.527	0.9667	3.06	29	0.005 *
	Cont.	11.4333	2.359	0.431				
Arithmetic	Exp.	9.5667	2.192	0.400	-0.6000	-1.81	29	0.080
	Cont.	10.1667	2.506	0.458				
Vocabulary	Exp.	11.3667	2.822	0.515	1.3667	5.34	29	0.001 *
	Cont.	10.0000	2.983	0.545				

* = significant (prob. $\leq$ .05)

Table 3 (continued)

Scale	Group	Mean	Standard Deviation	Standard Error	(Difference) Mean	t-Value	Degrees of Freedom	2-Tail Prob.
Comprehension	Exp.	11.6000	2.908	0.531	1.3333	5.05	29	0.000 *
	Cont.	10.2667	3.062	0.559				
Picture Completion	Exp.	10.9667	2.723	0.497	0.5333	1.25	29	0.220
	Cont.	10.4333	3.308	0.604				
Picture Arrangement	Exp.	11.7667	3.839	0.701	0.8000	3.53	29	0.001 *
	Cont.	10.9667	3.419	0.624				
Block Design	Exp.	11.7333	3.140	0.573	0.6000	2.26	29	0.031 *
	Cont.	11.1333	2.474	0.452				

* = significant (prob. ≤ .05)

Table 3 (continued)

Scale	Group	Mean	Standard Deviation	Standard Error	(Difference) Mean	t-Value	Degrees of Freedom	2-Tail Prob.
Object Assembly	Exp.	11.8333	3.108	0.567	0.0000	0.00	29	1.000
	Cont.	11.8333	3.992	0.729				
Coding	Exp.	13.0000	2.477	0.452	0.4667	1.18	29	0.247
	Cont.	12.5333	3.391	0.619				

* = significant (prob. ≦ .05)

Ho_1 was rejected for the following WISC-R subscales: Information, Similarities, Vocabulary, Comprehension, Picture Arrangement, and Block Design. Hypothesis One was retained for the following subscales: Arithmetic, Picture Completion, Object Assembly, and Coding.

The data in Table 3 are a comparison between the effects of feedback and no feedback regardless of student entry characteristics such as ability, age, and sex. They offer an interesting overall group comparison in terms of gains or loss. It can be seen that eight scales showed changes in the form of gains for the experimental group, five of which were found to be significant. Two scales showed losses for the experimental group, one of which was significant. A cursory inspection of these two scales shows them to be most visibly "school-related," one asking for recall of learned facts (Information), the other requiring computational mastery (Arithmetic).

Of those scales that showed learning gains for the experimental group, it seems that those encompassing an embedded "Gestalt," rule, or principle gained significantly, while those requiring a motor component, in part, did not show sufficient gains to make the experimental group perform significantly better than the control group. The subscales Similarities, Vocabulary, and Comprehension, for example, all showed significant gains for the experimental group. These scales have a built-in "understood" principle that carries throughout the task. In Similarities, the students are asked to find a common denominator in dissimilar objects. In Vocabulary, the task is to find a good synonym for progressively more difficult words. In Comprehension, the implied expectation is for the student to find a "socially accepted" answer to a problem. This characteristic of an underlying principle or rule is missing in those subscales that showed no significant difference. Picture Completion asks for discerning missing details in progressively more difficult cards that are dissimilar in perceptual make-up. Object Assembly requires the speedy assembly of puzzles, as such having a heavy motor component, as does Coding, also dependent on motor skills combined with speed of performance.

Coding deserves attention beyond the rest of the scales. It was the only scale not reinforced by feedback, due to the nature of the task that required rapid symbol reproduction within strictly defined time limits. In clinical interpretation, this scale is often given significance as a motivational indicator, coming at the end of a strenuous and anxiety-producing task. One might have speculated that the preceding feedback on the remaining nine scales might have affected the experimental group positively, according to behaviorist thinking, or negatively, according to the humanistic school of thought.

Inspecting the results of the groups as a whole, there was no motivational or deterring influence that could be detected.

HYPOTHESIS TWO

Ho_2 reads as follows: There is no difference between the experimental and control groups on the verbal and non-verbal summation scales. This hypothsis addresses itself to the factorial composition of the ten individual subscales by rearranging the data into two primary modes of cognition by which human abilities seem to express themselves, namely verbal and non-verbal learning. They also affix a norm-referenced quantification index to behavior, commonly known as verbal and non-verbal IQ. Table 4 summarizes the data to respond to Hypothesis Two.

Table 4

Comparison between the Experimental and Control Groups
on the WISC-R Verbal and Non-Verbal Summation Scales
(N = 30 for each group)

Scale	Group	Mean	Standard Deviation	Standard Error	(Difference) Mean	t-Value	Degrees of Freedom	2-Tail Prob.
Verbal Summation Scale	Exp.	106.5000	12.979	2.370	3.0667	3.24	29	0.003 *
	Cont.	103.4333	12.184	2.225				
Non-verbal Summation Scale	Exp.	112.9333	15.220	2.779	3.4333	2.37	29	0.025 *
	Cont.	109.5000	17.463	3.188				

* = significant (prob. ≤ .05)

The results permit the rejection of Ho_2 in both verbal and non-verbal performance. The experimental group, under the influence of feedback, significantly outperformed the control group in both instances. Feedback does make a difference both quantitatively and qualitatively in ways that can be measured and recorded factorially.

HYPOTHESIS THREE

Ho_3 reads as follows: There is no difference between the experimental group and the control group on the Full Scale Score. Table 5 offers a grand summary of the preceding data, taking the weight of the individual subscales as well as the verbal and non-verbal summation scales into account by an overall norm-referenced quantification, commonly known as "Full Scale IQ." Table 5 offers data in answer to Hypothsis Three.

Table 5

Comparison between the Experimental and Control Groups
on the WISC-R Full Scale Summation Scores
(N = 30 for each group)

Scale	Group	Mean	Standard Deviation	Standard Error	(Difference) Mean	*t*-Value	Degrees of Freedom	2-Tail Prob.
Full Summation Scale	Exp.	110.3333	13.914	2.540	3.4667	3.59	29	0.001 *
	Cont.	106.8667	14.890	2.718				

* = significant (prob. ≤ .05)

The statistically significant results permit the rejection of Ho_3. In overall performance, the experimental group scored significantly higher than the control group. Caution needs to be exercised, however, not to infer that the experimental group "gained in IQ." The experimental group gained in performance, but the scores are not equivalent IQ's since standardization rules, for the experimental group, were violated as dictated by the nature of the experiment.

In summary, Tables 3, 4, and 5 provided data in response to two parts of the three-part problem statement question. Feedback, given in response to lower order learning, seems to aid the mastery of higher order tasks in most instances but not all. This positive influence is strong enough to be reflected quantitatively and qualitatively in factorially discrete areas such as verbal and non-verbal performance as well as in overall performance as measured by norm-referenced results.

HYPOTHESES FOUR, FIVE, AND SIX

Tables 6 through 10 provide statistical data pertinent to the third part of the three-part question posed on page 89: Is there a significant interaction effect between feedback and student characteristics such as ability, age, and sex? This part of the statistical data is broken down as follows: Table 6 provides information pertaining to feedback affecting ability. Table 7 answers questions regarding feedback affecting age. Table 8 gives information regarding feedback and sex. All three tables are analyzed by means of a *t*-test analysis comparing the difference scores* of bright versus slow, young versus old, and male versus female students on main effects. Tables 9 and 10 complete the data by graphing significant two-way interaction effects, using an analysis of variance and a slightly reduced N of 28. Two cases (a "bright young male" and an "old slow female") were randomly chosen and removed from the sample in order to satisfy the prerequisite of θ equal Ns per cell as required by this particular statistical treatment.

Hypothesis Four (Interaction Effects)

Ho_4 reads as follows: The effects of feedback are constant with respect to ability. Table 6 relates to this question.

---

* difference score = experimental score minus control score for each N (D = E-C).

Table 6

Comparison of the Difference Scores (D = E-C) between Bright and Slow Children on all WISC-R Scales

Scale	Group	Mean	Standard Deviation	Standard Error	*t*-Value	Degrees of Freedom	2-Tail Prob.
Information	Bright	0.0667	1.668	0.431	2.61	27.21	0.015 *
	Slow	-1.4000	1.404	0.363			
Similarities	Bright	0.6000	1.957	0.505	-1.17	25.79	0.254
	Slow	1.3333	1.447	0.374			
Arithmetic	Bright	-0.8667	1.922	0.496	-0.80	27.65	0.430
	Slow	-0.3333	1.718	0.444			
Vocabulary	Bright	1.2667	1.335	0.345	-0.39	27.60	0.703
	Slow	1.4667	1.506	0.389			

* = significant (prob. ≤ .05)

Table 6 (continued)

Scale	Group	Mean	Standard Deviation	Standard Error	$\underline{t}$-Value	Degrees of Freedom	2-Tail Prob.
Comprehension	Bright	1.2667	1.486	0.384	-0.25	27.99	0.806
	Slow	1.4000	1.454	0.375			
Picture Completion	Bright	0.9333	2.219	0.573	0.94	27.74	0.356
	Slow	0.1333	2.446	0.631			
Picture Arrangement	Bright	1.0000	0.926	0.239	0.88	23.29	0.389
	Slow	0.6000	1.502	0.388			
Block Design	Bright	0.2667	1.100	0.284	-1.27	23.89	0.216
	Slow	0.9333	1.710	0.441			

* = significant (prob. $\leq$ .05)

Table 6 (continued)

Scale	Group	Mean	Standard Deviation	Standard Error	t-Value	Degrees of Freedom	2-Tail Prob.
Object Assembly	Bright	1.2667	1.944	0.502	3.25	27.19	0.003*
	Slow	-1.2667	2.314	0.597			
Coding	Bright	1.6000	2.324	0.600	3.34	21.32	0.003*
	Slow	-0.6667	1.234	0.319			
Verbal Summation Scale	Bright	3.0000	4.342	1.121	-0.07	25.36	0.945
	Slow	3.1333	6.069	1.567			
Non-verbal Summation Scale	Bright	7.2000	7.173	1.852	2.92	27.97	0.007*
	Slow	-0.3333	6.956	1.796			
Full Summation Scale	Bright	5.3333	4.530	1.170	2.03	27.05	0.052*
	Slow	1.6000	5.475	1.414			

* = significant (prob. ≤ .05)

Ho_4 was rejected for the following WISC-R scales: Information, Object Assembly, Coding, Non-Verbal Summation Scale, and Full Summation Scale. Ho_4 was retained for the remaining eight scales. Table 6 sheds additional light on the findings pertaining to Table 3. Looking at overall results (Table 3), one might conclude that feedback on the Information scale, for example, affected all children negatively. However, when breaking down the results according to ability, one can see that slower children were negatively affected, while more able children gained slightly by feedback. On Object Assembly, an even more interesting result was obtained: the substantial gain of the more capable group was cancelled out by an equally substantial loss for the less bright. Again, the Coding scale shows a sizeable gain for the bright group, while the less able group lost slightly. This differentiation carries into the Non-Verbal Summation Scale and the Full Summation Scale in behalf of the brighter children. Summarily, bright children gained substantially, while slow children suffered slightly on some of the scales that showed no overall difference when analyzed without respect to ability.

Although the differences between the different ability groups were not found to be significant on the remaining eight scales, it is interesting to inspect the results for directionality. The subscales Similarities, Vocabulary, and Comprehension, considered to be the most "cognitive" verbal scales, showed slow children gaining more than bright children, though both changes were in a positive direction. This arrangement is reversed for the non-verbal scales where more capable children do better than the slower ones on some but not all tasks due to feedback. In terms of overall gains, however, both bright and slow children gained verbally and on the Full Summation Scale score, while non-verbally the gain was clearly in behalf of the bright.

<u>Hypothesis Five (Interaction Effects)</u>

Ho_5 reads as follows: The effects of feedback are constant with respect to age. Table 7 is in response to this question.

Table 7

Comparison of the Difference Scores (D = E-C) between Young and Old Children on all WISC-R Scales

Scale	Group	Mean	Standard Deviation	Standard Error	t-Value	Degrees of Freedom	2-Tail Prob.
Information	Young	-0.6000	1.844	0.476	0.21	27.36	0.833
	Old	-0.7333	1.580	0.408			
Similarities	Young	0.4667	1.187	0.307	-1.63	22.34	0.118
	Old	1.4667	2.066	0.533			
Arithmetic	Young	-0.8000	1.897	0.490	-0.60	27.85	0.555
	Old	-0.4000	1.765	0.456			
Vocabulary	Young	0.7333	1.387	0.358	-2.74	26.94	0.011*
	Old	2.0000	1.134	0.293			

* = significant (prob. ≤ .05)

Table 7 (continued)

Scale	Group	Mean	Standard Deviation	Standard Error	t-Value	Degrees of Freedom	2-Tail Prob.
Comprehension	Young	1.2667	1.624	0.419	-0.25	26.70	0.806
	Old	1.4000	1.298	0.335			
Picture Completion	Young	0.6000	2.414	0.623	0.15	27.96	0.879
	Old	0.4667	2.326	0.601			
Picture Arrangement	Young	0.7333	0.961	0.248	-0.29	23.79	0.775
	Old	0.8667	1.506	0.389			
Block Design	Young	-0.4000	0.910	0.235	-5.19	26.27	0.000*
	Old	1.6000	1.183	0.306			

* = significant (prob. ≤ .05)

Table 7 (continued)

Scale	Group	Mean	Standard Deviation	Standard Error	t-Value	Degrees of Freedom	2-Tail Prob.
Object Assembly	Young	-0.2667	2.576	0.665	-0.59	27.87	0.562
	Old	0.2667	2.404	0.621			
Coding	Young	0.2000	2.484	0.641	-0.67	25.74	0.509
	Old	0.7333	1.831	0.473			
Verbal Summation Scale	Young	1.2667	4.773	1.232	-2.00	27.88	0.053 *
	Old	4.8667	5.097	1.316			
Non-verbal Summation Scale	Young	1.3333	8.958	2.313	-1.48	25.27	0.151
	Old	5.5333	6.368	1.644			
Full Summation Scale	Young	1.3333	5.802	1.498	-2.38	24.19	0.026 *
	Old	5.6000	3.814	0.985			

* = significant (prob. ≤ .05)

Ho_5 was rejected for the following WISC-R scales: Vocabulary, Block Design, Verbal Summation Scale, and Full Summation Scale. Ho_5 was retained for the remaining nine scales. Older children gained more by feedback than young children on the two scales considered strong scholastic predictors in both the verbal and non-verbal scales (Vocabulary and Block Design), and these gains were strong enough to carry over into the Verbal and Full Summation scales. In view of the fact that age was not a significant factor in the remaining nine subscales, an alternate explanation is possible. Length of task rather than age may have influenced these results. The Vocabulary subscale is by far the longest in numbers of items (32); Block Design is fairly long. Older children may have been given proportionately more feedback than younger children before they reached their ceiling on these particular scales. Older children were also favored on the remaining non-significant scales in Similarities, Vocabulary, and Comprehension (all verbal scales) while non-verbally the difference is small.

Hypothesis Six (Interaction Effects)

Ho_6 reads as follows: The effects of feedback are constant with respect to sex. Table 8 is in response to this question:

Table 8

Comparison of the Difference Scores (D = E-C) between Males and Females on all WISC-R Scales

Scale	Group	Mean	Standard Deviation	Standard Error	$\underline{t}$-Value	Degrees of Freedom	2-Tail Prob.
Information	Male	-1.0000	1.620	0.393	-1.24	24.96	0.228
	Female	-0.2308	1.739	0.482			
Similarities	Male	0.5294	1.375	0.333	-1.54	20.07	0.138
	Female	1.5385	2.025	0.562			
Arithmetic	Male	-0.5882	2.210	0.536	0.04	25.57	0.966
	Female	-0.6154	1.193	0.331			
Vocabulary	Male	1.5882	1.372	0.333	0.98	25.29	0.335
	Female	1.0769	1.441	0.400			

* = significant (prob. ≤ .05)

Table 8 (continued)

Scale	Group	Mean	Standard Deviation	Standard Error	t-Value	Degrees of Freedom	2-Tail Prob.
Comprehension	Male	1.2353	1.522	0.369	-0.42	27.03	0.675
	Female	1.4615	1.391	0.386			
Picture Completion	Male	1.0000	2.151	0.522	1.24	23.73	0.226
	Female	-0.0769	2.499	0.693			
Picture Arrangement	Male	0.4118	1.278	0.310	-2.12	27.89	0.043*
	Female	1.3077	1.032	0.286			
Block Design	Male	0.3529	1.412	0.342	-1.06	25.14	0.300
	Female	0.9231	1.498	0.415			

* = significant (prob. ≤ .05)

Table 8 (continued)

Scale	Group	Mean	Standard Deviation	Standard Error	$\underline{t}$-Value	Degrees of Freedom	2-Tail Prob.
Object Assembly	Male	-0.1765	2.675	0.649	-0.45	27.71	0.654
	Female	0.2308	2.242	0.622			
Coding	Male	0.4706	2.348	0.570	0.01	27.67	0.991
	Female	0.4615	1.984	0.550			
Verbal Summation Scale	Male	2.3529	4.885	1.185	-0.84	23.91	0.408
	Female	1.0000	5.612	1.557			
Non-verbal Summation Scale	Male	2.8824	9.158	2.221	-0.45	27.72	0.655
	Female	4.1538	6.256	1.735			
Full Summation Scale	Male	2.7059	6.162	1.495	-0.95	27.17	0.349
	Female	4.4615	3.886	1.078			

* = significant (prob. ≤ .05)

Hypothesis Six was retained for all except the Picture Arrangement subscale. On this particular task, females scored higher than males under the influence of feedback, though both gains were in a positive direction. This particular scale assesses, in part, a child's acuity in sizing up a social situation and ordering events according to time. Perhaps there is a culturally determined factor at play here, although this must be conjecture in light of the overall non-significant results pertaining to feedback in interaction with sex.

Two-Way and Three-Way Interaction Effects

An analysis of variance was performed on all difference scores to test for two-way and three-way interaction effects between feedback and ability, age, and sex. The complete statistical results are listed in Appendix D. No significant differences were found except in one instance between ability and sex on the Object Assembly scale, a result that carried over into the Non-Verbal Summation Scale. The complete analysis of variance tables for these two scales as well as a visual representation of breakdowns according to cells are offered below.

Table 9

Analysis of Variance for the WISC-R Difference Scores on the Subscale Object Assembly

Source of Variation	Sum of Squares	DF	Mean Square	F	Signif. of F
Main Effects	45.062	3	15.021	4.458	0.015
Sex	0.916	1	0.916	0.272	0.608
YO	2.791	1	2.791	0.828	0.374
BS	40.844	1	40.844	12.123	0.002*
2-Way Interactions	47.105	3	15.702	4.660	0.013
Sex YO	0.045	1	0.045	0.013	0.909
Sex BS	34.133	1	34.133	10.131	0.005*
YO BS	10.761	1	10.761	3.194	0.089
3-Way Interactions	8.450	1	8.450	2.508	0.129
Sex YO BS	8.450	1	8.450	2.508	0.129
Explained	100.617	7	14.374	4.266	0.005
Residual	67.383	20	3.369		
Total	168.000	27	6.222		

N = 28 * = significant (prob. ≤ .05)

Table 10

Analysis of Variance for the WISC-R Difference Scores on the Performance Summation Scale

Source of Variation	Sum of Squares	DF	Mean Square	F	Signif. of F
Main Effects	576.648	3	192.216	3.952	0.023
Sex	4.252	1	4.252	0.087	0.771
YO	177.408	1	177.408	3.648	0.071
BS	441.752	1	441.752	9.084	0.007*
2-Way Interactions	198.483	3	66.161	1.360	0.283
Sex YO	9.557	1	9.557	0.197	0.662
Sex BS	197.633	1	197.633	4.064	0.053*
YO BS	0.418	1	0.418	0.009	0.927
3-Way Interactions	39.200	1	39.200	0.806	0.380
Sex YO BS	39.200	1	39.200	0.806	0.380
Explained	814.331	7	116.333	2.392	0.060
Residual	972.633	20	48.632		
Total	1786.964	27	66.184		

N = 28   * = significant (prob. $\leq$ .05)

Table 11

Interaction Effects of Feedback on Sex and Ability on the Subscale Object Assembly

	Slow	Bright	
Male	$\overline{X} = -2.222$ N = 8	$\overline{X} = 2.143$ N = 8	16
Female	$\overline{X} = .600$ N = 6	$\overline{X} = .286$ N = 6	12
	14	14	

Table 12

Interaction Effects of Feedback on Sex and Ability on the WISC-R Non-Verbal Summation Scale

	Slow	Bright	
Male	$\overline{X} = -2.555$ N = 8	$\overline{X} = 9.42$ N = 8	16
Female	$\overline{X} = 4.000$ N = 6	$\overline{X} = 5.142$ N = 6	12
	14	14	

Slow males were negatively affected by feedback on Object Assembly, while bright males and all females gained in performance when given feedback. It should be mentioned here that this particular scale is considered one of the weaker cognitive WISC-R scales in that the task itself is too short (only 4 items) to give an adequate performance sample. It is also one of the more visibly "manipulative" tasks requiring strong dexterity in a puzzle assembly of items such as a racing car, for example. One might speculate that slow males may have interpreted the task as a "prestige test" and may have felt unduly anxious.

SUMMARY

Chapter Four offered statistical evidence relevant to the six null hypotheses comprising this study. Within the boundaries set by the nature of the WISC-R scales, all six hypotheses could be rejected. The magnitude and directionality of learning under the influence of feedback varied according to the nature of the task. Ability, age, and sex as student variables also seemed to be influenced differentially by feedback, with ability showing the strongest variation in behalf of the bright, age showing some variation favoring older children, and sex showing very little change. In terms of overall gains, the evidence is in support of feedback as a means of instructional quality control.

Chapter 5

CONCLUSIONS AND RECOMMENDATIONS

INTRODUCTION

The intent of this study was to measure as closely as possible the magnitude of overall learning under the influence of feedback, as well as the magnitude of the distance between boys and girls of different abilities and ages. In terms of overall learning, this was done by a scale score comparison of 13 WISC-R scales (Tables 3 through 5) between the experimental and control groups, on the assumption that performance on scales comprising relatively discrete, hierarchically ordered tasks was representative of overall cognitive performance. Only to the degree that this assumption is true can inferences to classroom application be made. However, this study relied on the triple strength of 1) a sample of 30 chronologically spaced sets of twins with built-in controls as to ability, age, and sex, 2) a strictly controlled behavioral setting, and 3) the advantage of translation of behavioral raw scores into refined, norm-referenced quantification. The obtained results may therefore be taken as relatively "pure" samples of learning that did or did not occur under the influence of informational feedback.

In terms of specific differences between children of different abilities and ages, an analysis was made by a *t*-test and analysis of variance comparison between the difference scores of rank-ordered groups according to IQ and chronological age (Tables 6 through 10). The organismic variables of "bright-slow" and "young-old" were arbitrarily derived by using the median as the point of division into two respective groups for both age and ability. The limitations inherent in this arbitrary division were acknowledged but seen as being outweighted by the advantages of data derived from a twin sample in this particular experiment. Sex was a natural, discrete variable that lent itself to a partition of groups of boys and girls.

Chapter 5 summarizes important points made in Chapters 1 through 4 and concludes with implications and recommendations based on the findings of this study. For additional clarity, Chapters 1 through 4 are briefly reviewed.

Chapter 1 pointed out that traditionally, school administrators have treated scholastic feedback as an absolute--as something that "ought" to be given or withheld, according to prevailing ideological beliefs. Instead of a political or pedagogical "either-or" approach pertaining to feedback, Chapter 1 suggested an alternative approach based on the judicious use of feedback in light of certain characteristics of learning, and in light of important learner variables such as ability, age, and sex. A need was hereby established for an empirical analysis of certain aspects of learning regarding both learner and task.

Chapter 2 summarized and highlighted both theoretical and practical aspects of feedback as pertaining to hierarchical learning, and as pertaining to current instructional practices embodying feedback, such as operant conditioning and Learning-For-Mastery approaches. Special attention was given to three advocates of learning: Robert Gagne, David Ausubel, and Jerome Bruner, and their respective theoretical positions regarding feedback. A variety of tangential research pertaining to feedback and hierarchical learning, especially in relation to ability, age, and sex, was also cited as evidence that a sharper empirical focus was needed.

Chapter 3 translated such a focus into experimental terms, describing a controlled learning setting that would address itself to three interrelated questions seen as central to this study. These questions were condensed and paraphrased as follows: 1) Will feedback in response to lower order learning facilitate the mastery of higher order tasks? 2) If so, in what magnitude qualitatively and quantitatively, as measured by relatively concise scale score results? 3) Will student characteristics such as ability, age, and sex interact with feedback in a significant way? A strictly controlled learning setting, embodying progressively more difficult tasks subject to precise experimental control, with informational feedback as the independent variable and various psychometric results as the dependent variables, was seen as providing useful and substantive answers to the six hypotheses posed in this study.

Chapter 4 presented the statistical results and narrative summaries of the empirical findings. Several overall patterns could be discerned. Informational feedback proved to be a positive influence on learning in ways that were both task-specific and child characteristics-specific. The subjects showed themselves to be responsive to informational feedback in differing degrees and directions. The student characteristics of ability, age, and sex further clarified the influence of feedback by showing that feedback, though of benefit overall, aided certain subgroups more than others, and seemed to deter learning in some.

IMPLICATIONS

Qualitatively and quantitatively, higher order learning did occur in response to lower order reinforcement in 50 percent of all tasks, as measured by overall group performance. Feedback seemed to deter performance in 10 percent of the tasks, and had no seeming influence in 40 percent of the tasks. This is evidence consistent with the use of open, objective feedback in the form of grading, homogeneous grouping, merit systems and other means intended to make performance "pay off." This position has long been advocated by writers such as Ebel and Popham, and has been incorporated in part by instructional strategies such as Bloom's Learning-for-Mastery approach. The overall evidence obtained in this study would seem to weaken the positions taken by Green, Glasser, and others who hold that painful feelings of failure and exposure caused by negative feedback depress learning in children to a degree that makes these practices harmful for all.

When viewing the directionality of the results, one can detect that the two scales where performance was negatively affected in comparison to the eight where performance was in a positive direction were scales quite easily identifiable as tapping school-related learning. Both recall of facts and computational skill are everyday tasks children are asked to perform in the classroom. It is conceivable that a certain learner anxiety has become attached to certain performance tasks, and that this emotional anxiety might override potential benefits of feedback for some children in some cases, at least. While feedback tends to be beneficial overall, there may be occasions where an exception may be made in favor of withholding of feedback. Just what these occasions might be would be for further research to sort out.

In terms of <u>specific</u> performance of certain subgroups divided according to ability, age, and sex, the data would seem to complement and refine the information gained by an overall scale score comparison. Of strongest significance are the data pertaining to ability. In almost all instances, high ability children seemed to benefit from feedback, while children of lesser ability benefited generally but not always. In other words, the motivational loss often assumed to be the result of conventional feedback devices pertains only to children with somewhat limited ability, with some exceptions. Administrators will benefit from the knowledge that a "price tag" in terms of learning is attached especially to the performance of scholastically capable children where informational feedback is withheld.

Some scales showed no statistical significance between the distance of the ability groups, but both groups gained by being given informational feedback. These findings are congruent with previous research that showed an increase in distance in some cases and a narrowing in distance in others in learning tasks that utilize feedback. Two of these strategies are the previously mentioned Learning-for-Mastery approach or the various operant conditioning approaches. Future years might see more effort directed at instructional strategies that refine and incorporate informational feedback by taking ability levels into account.

When comparing younger versus older children by means of statistical treatment of difference scores, a similar but slightly less clear picture emerges. Again, as measured by overall qualitative and quantitative results, both groups gained by feedback, but certain tasks seemed to carry a motor component that worked in behalf of the older group but not in behalf of the younger.

For example, in four out of five non-verbal scales, the younger children did less well than the older group, although only two were statistically significant. One alternate explanation offers itself: it is quite possible that the amount of feedback (more reinforced items for the old, due to the nature of the WISC-R ceilings) rather than age itself made the difference. It is also possible that older children are more cued in to the benefits of feedback due to prior learning. Homogeneous grouping, "tracking," various curricular stratification plans, etc., tend to be emphasized more as children progress chronologically, with non-grading "open classroom" practices happening more frequently in lower grades. The obtained data pertaining to age groups would seem to give partial support to stepped-up feedback practices in later years.

The nature of sex-typed behavior is often held to be responsible for differential performance in schools, the conventional argument usually being that boys seek feedback more vigorously since both home and school reinforce the image of a competitive male.

The obtained data of this study pertaining to feedback and sex do not bear this out. On the whole, feedback seemed to influence learning in ways that were not sex-related except in isolated instances.

Another way of looking at the results is in terms of leading theorists of hierarchical learning and their stance with respect to feedback. Robert Gagne, for example, advocates the systematic use of feedback by making a distinction between concrete (lower-order) and abstract (higher order) tasks, advocating strong operant conditioning principles in the beginnings of learning and relying more upon "internalized" reinforcement in higher-order tasks. His theory was supported on some WISC-R scales but not all. The deciding factor seemed to be the degree of continuity within a scale versus scales comprised of items that were relatively discrete, such as recall of independent facts. Also, lower order feedback did not help with items that might be labelled school-related, as, for example, computational skills. David Ausubel's theoretical formulation emphasizing verbal learning, and concurring with Gagne on some but not all premises regarding the value of reinforcement, seemed to gain some empirical support by the results of this study. Verbal learning seemed to be aided more readily and visibly by feedback than non-verbal learning, especially for older children and for the bright. The results of this study overall do not support Jerome Bruner's contention that non-reinforced exploration in and of itself will carry hierarchical learning. Although some WISC-R scales encourage exploration and manipulation of tasks, in no instance was the performance of the non-reinforced control group superior to the experimental group who were given step-by-step feedback. Age seemed to be a factor in differential learning, but not as crassly obvious as often assumed and stressed. Only 20 percent of the subscales were found to differentiate the performance of young in contrast to older children, although this contrast was strong enough to influence the summation scales. Developmental educators who support non-grading, non-evaluative practices in lower grades may conceivably use the data of this study in their behalf, although it should be pointed out that for the majority of the tasks, feedback was of benefit for all children regardless of chronological age.

In summary, it may be said that this study showed feedback to be more beneficial in verbal than non-verbal learning, more beneficial for more scholastically capable children and for the older group, but that in absolute terms, as measured by the degree as well as the direction of learning, the presence or absence of feedback determined the magnitude in measurable terms, a finding congruent with most research in the field.

RECOMMENDATIONS

Further empirical explorations might take the following direction:

1) A descriptive study that would attempt to "quantify" the concept of feedback as it is found in a) administrator's awareness, b) in the literature, and c) in academic courses taught. Compiling the background research for this study showed that the concept itself is rather invisible. The term can be found in auxiliary realms such as systems theory, cybernetics theory, or communications theory. Feedback is not seen or treated as a concept in and of itself as having properties for educational-instructional quality control. Many if not most texts in school administration that have been published within the last twenty years do not yet carry the term in their indices.

2) A longitudinal replication of this study in order to measure direct as well as long-term effects of feedback on hierarchical learning. A test-retest situation measuring feedback on reinforced items by using parallel measuring devices might shed additional light on the intricacies of hierarchical learning in light of age, ability, and sex.

3) The effects of feedback in light of ability and age by aligning feedback more closely to conventional classroom practices, and using a larger sample and a more classroom-like setting. The

present study was narrowly circumscribed by the restrictions inherent in the sample, the task, and the instrument. Real-life learning is much more a function of environmental influences that cannot be discerned and isolated, much less controlled to the degree that this study permitted. A more dynamic, flux-like setting such as a contemporary classroom provides might make sharp distinctions less visible.

BIBLIOGRAPHY

BOOKS

Arlin, M. N. *Contemporary Issues in Educational Testing*, ed. H. F. Crombag and D. N. DeGruiter. Paris: Mouton, 1974.

Ausubel, D. P. *Educational Psychology: A Cognitive View.* New York: Holt, Rinehart and Winston, 1968.

__________. *The Psychology of Meaningful Verbal Learning.* New York: Grune and Stratton, 1963.

Bennis, W. G., K. D. Benne, R. Chin, and K. E. Corey, *The Planning of Change.* New York: Holt, Rinehart and Winston, 1976.

Berry, G. *PSI: 41 Germinal Papers*, ed. J. G. Sherman. Menlo Park: W. A. Benjamin, 1974.

Block, J. H., and L. W. Anderson. *Mastery Learning in Classroom Instruction.* New York: Macmillan, 1975.

__________ and R. B. Burns. *Review of Research in Education*, ed. L. S. Shulman. Itaska: Peacock, 1976.

Bloom, B. S. *Confronting Curriculum Reform*, ed. E. W. Eisner. Boston Little, Brown, 1971.

__________. *Schools, Society, and Mastery Learning*, ed. J. H. Block. New York: Holt, Rinehart and Winston, 1974.

Broderick, C., and J. Bernard, eds. *The Individual, Sex, and Society: A SIECUS Handbook for Teachers and Counselors.* Baltimore: The Johns Hopkins Press, 1969.

Bruner, J. *The Process of Education.* Boston: Harvard University Press, 1962.

__________. Studies in Cognitive Growth. New York: Wiley, 1968.

__________, J. J. Goodnow, and G. A. Austin. A Study of Thinking. New York: Wiley, 1956.

Buros, O. K., ed. The Seventh Mental Measurement Yearbook. Highland Park: The Gryphon Press, 1972.

Campbell, D. T., and J. C. Stanley. Handbook of Research on Teaching, ed. N. L. Gage. Chicago: Rand McNally, 1963.

Cole, C., S. Martin, and J. Vincent. Behavior Research and Technology in Higher Education, ed. J. M. Johnston. Springfield, Ill.: Charles C. Thomas, 1975.

Crowder, N. A. Automatic Teaching: The State of the Art, ed. E. Galanter. Wiley, 1959.

__________. Teaching Machines and Programmed Learning: A Source Book, ed. A. A. Lumsdaine and R. Glaser. Washington, D. C.: National Education Association, 1960.

Dale, E. Programmed Instruction, ed. P. C. Lange. University of Chicago Press, 66th Yearbook, NSSE, 1967.

Davis, M. L. Behavior Research and Technology in Higher Education, ed. J. M. Johnston. Springfield, Ill.: Charles C. Thomas, 1975.

Evans, J. L. Teaching Machines and Programmed Learning. II: Data and Directions, ed. R. Glaser. Washington, D. C.: National Education Association, 1965.

Fowler, H. Curiosity and Exploratory Behavior. New York: Macmillan, 1965.

Gagne, R. M. The Conditions of Learning. New York: Rinehart and Winston, 1965.

__________. Perspectives of Curriculum Evaluation. New York: Rand McNally, 1967.

Getzels, J. W., and J. T. Dillon. Second Handbook of Research on Teaching, ed. R. W. Travers. Chicago: Rand McNally, 1973.

Glasser, A. J., and I. L. Zimmerman. Clinical Interpretation of the Wechsler Intelligence Scale for Children. New York: Grune and Stratton, 1967.

Glasser, W. Schools Without Failure. New York: Harper and Row, 1969.

Holland, J. G. Automatic Teaching: The State of the Art, ed. E. Galanter. Wiley, 1959.

__________. Teaching Machines and Programmed Learning. II: Data and Directions, ed. R. Glaser. Washington, D. C.: National Education Association, 1965.

Homme, L. How to Use Contingency Contracting in the Classroom. Champaign, Ill.: Research Press, 1969.

Jensen, A. R. Genetics and Education. New York: Harper and Row, 1972.

Kneller, G. F. Foundations of Education. New York: John Wiley and Sons, 1971.

Lawther, J. D. The Learning of Physical Skills. Englewood Cliffs, N. J.: Prentice-Hall, 1968.

Lumsdaine, A. A. Handbook of Research on Teaching, ed. N. L. Gage. Chicago: Rand McNally, 1963.

__________ and R. Glaser, eds. Teaching Machines and Programmed Learning: A Source Book. National Education Association, 1960.

Meyer, S. R. Automatic Teaching: The State of the Art, ed. E. Galanter. Wiley, 1959.

Miller, N. E. Psychopathology Today: Experimentation, Theory, and Research, ed. W. S. Sahakian. Itaska, Ill.: Peacock, 1970.

Nixon, J. E., and L. F. Locke. Second Handbook of Research in Teaching, ed. R. M. Travers, American Educational Research Association. Chicago: Rand McNally, 1973.

Nuthall, G., and I. Snook. Second Handbook of Research in Education, ed. R. M. Travers. Chicago: Rand McNally, 1970.

Piaget, J. The Science of Education and the Psychology of the Child. New York: Grossman, 1970.

Popham, W. J. Educational Evaluation. Englewood Cliffs, Prentice Hall, 1975.

Porter, D. Automatic Teaching: The State of the Art, ed. E. Galanter. Wiley, 1959.

Rosati, P. Behavior Research and Technology in Higher Education, ed. J. M. Johnson. Springfield: Charles C. Thomas, 1975.

Schramm, W. L. The Research on Programmed Instruction. Washington, D. C.: U.S. Government Printing Office, 1964.

Scriven, M. Problems and Prospects for Individualized Education. Berkeley: McCutchan, 1975.

Skinner, B. F. Science and Human Behavior. New York: MacMillan, 1953.

__________. The Technology of Teaching. New York: Appleton-Century-Crofts, 1968.

Thorndike, E. L. The Psychology of Learning. New York: Columbia University, 1913.

Wechsler, D. Manual for the Wechsler Intelligence Scale for Children - Revised. New York: The Psychological Corporation, 1974.

Wolf, M. M., and T. R. Risley. The Nature of Reinforcement, ed. R. Glaser. New York: Academic Press, 1971.

ARTICLES

Ammons, R. B. "Effects of Knowledge of Performance: A Survey and Tentative Theoretical Formulation," Journal of General Psychology, 54 (1956), 279-299.

Anderson, R. C., R. W. Kulhavy, and T. Andre. "Feedback Procedures in Programmed Instruction," Journal of Educational Psychology, 62 (1971), 148-56.

Annett, J., and H. Kay. "Knowledge of Results and 'Skilled Performance'," Occupational Psychology, 31 (1957), 69-79.

Ausubel, D. P. "Cognitive Theory of School Learning," Psychology in Schools, 6 (October, 1969), 331-5.

__________. "The Use of Advance Organizers in the Learning and Retention of Meaningful Verbal Material," Journal of Educational Psychology, 51 (1960), 267-72.

Baratz, J. C., and J. H. Moskowitz. "Proposition 13: How and Why It Happened," Phi Delta Kappan, 60 (September, 1978), 9-11.

Battle, E. W., and B. Lacey. "A Context for Hyperactivity in Children over Time," Child Development, 43 (1972), 757-73.

Bem, S. L. "Verbal Self-control: The Establishment of Effective Self-instruction," Journal of Experimental Psychology, 74 (1967), 485-91.

Bhavnani, R. M., and C. Hutt. "Divergent Thinking in Boys and Girls," Journal of Child Psychology and Psychiatry, 13 (1972), 121-27.

Biber, H., L. B. Miller, and J. L. Dyer. "Feminization in Preschool," Developmental Psychology, 7 (1972), 86.

Bilodeau, E. A., and I. M. Bilodeau. "Motor Skills Learning," Annual Review of Psychology, 12 (1961), 243-280.

Block, J. H., and M. Tierney. "An Exploration of Two Correctional Procedures Used in Mastery Learning Approaches to Instruction," Journal of Educational Psychology, 66 (November, 1974), 962-67.

Born, D. G., S. M. Gledhill, and M. L. Davis. "Examination Performance in Lecture-Discussion and Personalized Instruction Courses," Journal of Applied Behavioral Analysis, 5 (1972), 33-43.

Brown, J. L. "Effects of Logical and Scrambled Sequences in Mathematical Materials on Learning with Programmed Instruction Materials," Journal of Educational Psychology, 61 (1970), 41-45.

Bruner, J. S. "The Act of Discovery," Harvard Educational Review, 31 (1961), 21-32.

__________. "The Skill of Relevance and the Relevance of Skills," Saturday Review, April 18, 1970, pp. 66-8, 78-9.

Burt, C. "The Genetic Determination of Monozygotic Twins Reared Together and Apart," British Journal of Psychology, 57 (1966), 137-53.

Clifford, M. M., and T. A. Cleary. "The Relationship between Children's Academic Performance and Achievement Accountability," Child Development, 43 (1972), 647-55.

Coons, C. C. "Non-Promotion: A Dead End Road," Phi Delta Kappan, 58 (May, 1977), 701-2.

Costin, F. "Lecturing Versus Other Methods of Teaching: A Review of Research," British Journal of Educational Technology, 3 (1972), 4-30.

Dweck, C. S., and N. D. Repucci. "Learned Helplessness and Reinforcement Responsibility in Children," Journal of Personal Social Psychology, 25 (1973), 109-16.

Ebel, R. L. "The Case for Minimum Competency Testing," Phi Delta Kappan, 59 (April, 1978), 547.

__________. "Educational Tests: Valid? Biased? Useful?" Phi Delta Kappan, 56 (October, 1975), 83-88.

Eisenberg, T. A., and H. H. Walbesser. "Learning Hierarchies--Numerical Considerations," Journal for Research in Mathematics Education, 2 (1971), 244-56.

Enochs, J. C. "Modesto, California: A Return to the Four R's," Phi Delta Kappan, 59 (May, 1978), 609-10.

Erlenmeyer-Kimling, L., and L. F. Jarvik, "Genetics and Intelligence: A Review," Science, 142 (1963), 1477.

Farmer, J., G. Lachter, J. Blaustein, and D. Cole. "The Role of Proctoring in Personalized Instruction," Journal of Applied Behavioral Analysis, 5 (December, 1972), 401-4.

Fletcher, J. D., and R. C. Atkinson. "Evaluation of the Stanford CAI Program in Initial Reading," Journal of Educational Psychology, 63 (1972), 597-602.

Frase, L. T. "Cybernetic Control of Memory While Reading Connected Discourse," Journal of Educational Psychology, 60 (1969), 49-55.

__________. "Paragraph Organization of Written Materials: The Influence of Conceptual Clustering upon the Level and Organization of Recall," Journal of Educational Psychology, 60 (1969), 394-401.

__________. "Boundary Conditions for Mathemagenic Behaviors," Review of Educational Research, 40 (1970), 337-47.

Friedmann, M. M. "Spanish-Bilingual Students and Intelligence Testing," Thrust, 3 (November, 1973), 20-23.

Gagne, R. M. "Learning Hierarchies," Educational Psychologist, 6 (1968), 1-9.

__________. "Domains of Learning," Interchange, 3 (1972), 1-8.

__________. "Factors in Acquiring Knowledge of a Mathematical Task," Psychological Monographs, 76 (1962), No. 526.

__________, and L. T. Brown. "Some Factors in the Programming of Conceptual Learning," Journal of Experimental Psychology, 62 (1961), 313-21.

__________, and W. D. Rohwer, Jr. "Instructional Psychology," Annual Review of Psychology, 20 (1969), 381-418.

Gibson, E. J. "Learning to Read," Science, 148 (1965), 1066-72.

Gjesme, T. "Achievement-related Motives and School Performance for Girls," Journal of Personal Social Psychology, 26 (1973), 131-36.

Green, R. L. "Tips on Educational Testing: What Teachers and Parents Should Know," Phi Delta Kappan, 56 (October, 1975), 89-93.

Hartley, J. "Programmed Instruction 1954-1974: A Review," Programmed Learning and Educational Technology, 11 (1974), 278-91.

Holt, J. "I Oppose Testing, Marking, and Grading," Today's Education, 60 (March, 1971), 28-31.

Howell, M. L. "Use of Forcetime Graphs for Performance Analysis in Facilitating Motor Learning," Research Quarterly, 27 (1956), 12-22.

Kogan, N., and E. Pankove. "Creative Ability over a Five-year Span," Child Development, 43 (1972), 427-42.

Kulhavy, R. W., and R. C. Anderson, "Delay-retention Effect with Multiple-choice Tests," Journal of Educational Psychology, 63 (1972), 505-12.

Lee, P. C., and A. L. Wolinsky. "Male Teachers of Young Children: A Preliminary Empirical Study," Young Children, 28 (1973), 342-52.

Levitin, T. E., and J. D. Chanani. "Responses of Female Primary School Teachers to Sex Typed Behaviors in Male and Female Children," Child Development, 43 (1972), 1309-16.

Levy, P., and J. Tucker, "Differential Effects of Streaming on Primary School Attainment," British Journal of Educational Psychology, 42 (1972), 75-9.

Lloyd, K., and N. Knutzen. "A Self-paced Programmed Undergraduate Course in the Experimental Analysis of Behavior," Journal of Applied Behavior Analysis, 2 (1969), 125-33.

Malina, R. M. "Effects of Varied Information Feedback Practice Conditions on Throwing Speed and Accuracy," Research Quarterly, 40 (1969), 135-45.

Maxwell, A. E. "A Factor Analysis of the Wechsler Intelligence Scale for Children," British Journal of Educational Psychology, 29 (1959), 237-41.

McCall, R. B., P. S. Hogarty, and N. Hurlburt. "Transition in Infant Sensorimotor Development and the Prediction of Childhood IQ," American Psychologist, 27 (1972), 728-48.

McKeachie, W. J. "Instructional Psychology," Annual Review of Psychology, 25 (1974), 161-66.

__________. "The Decline and Fall of the Laws of Learning," Educational Researcher, 3 (July, 1974), 7-11.

McMichael, J. S., and J. R. Corey. "Contingency Management in an Introductory Psychology Course Produces Better Learning," Journal of Applied Behavioral Analysis, 2 (1969), 79-83.

Merrill, M. D. "Necessary Psychological Conditions for Defining Instructional Outcomes," Educational Technology, 2 (1971), 34-9.

__________, K. Barton, and L. E. Wood. "Specific Review in Learning a Hierarchical Imaginary Science," Journal of Educational Psychology, 61 (1970), 102-109.

Miller, N. E. "Biofeedback and Visceral Learning," Annual Review of Psychology, 29 (1978), 373-404.

__________. "Extending the Domain of Learning," Science, 152 (1966), 676.

__________. "Learning of Visceral and Glandular Responses," Science, 163 (1969), 434-45.

More, A. J. "Delay of Feedback and the Acquisition and Retention of Verbal Materials in the Classroom," Journal of Educational Psychology, 60 (1969), 339-42.

Nazarro, J. R., J. C. Todorov, and J. N. Nazarro. "Student Ability and Individualized Instruction," Journal of College Science Teaching, 2 (1972), 29-30.

Niedermeyer, F., J. Brown, and B. Sulzen. "Learning and Varying Sequences of Ninth-grade Mathematics Materials," Journal of Experimental Education, 37 (1969), 61-66.

Olson, G. H. "A Multivariate Examination of the Effects of Behavioral Objectives, Knowledge of Results, and the Assignment of Grades on the Facilitation of Classroom Learning," PhD thesis, Dissertation Abstracts International, 32 (1972), 6214-15.

Oner, N. P. "Impact of Teacher Behavior and Teaching Technique on Learning by Anxious Children," Dissertation Abstracts International, 32 (1972), 6215.

Payne, D. A., D. R. Krathwohl, and J. Gordon. "The Effects of Sequence on Programmed Instruction," American Educational Research Journal, 4 (1967), 125-32.

Popham, W. J. "Curriculum Materials," Review of Educational Research, 39 (1969), 319-38.

Pressey, S. L. "A Simple Apparatus Which Gives Tests and Scores--and Teaches," School and Society, 23 (1926), 373-6.

Robinson, D. W. "Book Review of Shirley Boes Neill's The Competency Movement," Phi Delta Kappan, 59 (May, 1978), 639.

Rohwer, W. D., Jr., and J. R. Levin. "Action, Meaning, and Stimulus Selection in Paired Associate Learning," Journal of Verbal Learning and Verbal Behavior, 7 (1968), 137-41.

Rothkopf, E. "The Concept of Mathemagenic Activities," Review of Educational Research, 40 (1970), 325-36.

Saretzky, G. "The Strangely Significant Case of Peter Doe," Phi Delta Kappan, 54 (May, 1973), 589-92.

Sassenrath, J. M. "Effects of Delay of Feedback and Length of Postfeedback Interval on Retention of Prose Material," Psychology in Schools, 9 (1972), 194-97.

__________ and G. D. Yonge, "Effects of Delayed Information Feedback and Feedback Cues in Learning on Delayed Retention," Journal of Educational Psychology, 60 (1969), 174-77.

Sheppard, W. C., and H. G. MacDermot. "Design and Evaluation of a Programmed Course in Introductory Psychology," Journal of Applied Behavior Analysis, 3 (1970), 5-11.

Sidman, M., and L. T. Stoddard. "The Effectiveness of Fading in Programming a Simultaneous Form Discrimination for Retarded Children," Journal of Experimental Analysis of Behavior, 10 (1967), 3-15.

Skinner, B. F. "The Science of Learning and the Art of Teaching," Harvard Educational Review, 24 (1954), 86-97.

__________. "The Free and Happy Student," Phi Delta Kappan, 55 (September, 1973), 13-16.

Stolurow, L. M. "Teaching by Machine," Cooperative Research Monographs, 6 (1961), 14-20.

__________. "A Model and Cybernetic System for Research on the Teaching-Learning Process," Programmed Learning, 2 (1965-b), 138-57.

Sturges, P. T. "Information Delay and Retention: Effect of Information in Feedback and Tests," Journal of Educational Psychology, 63 (1972), 32-43.

Tennyson, R. D., and M. D. Merrill. "Hierarchical Models in the Development of a Theory of Instruction: A Comparison of Bloom, Gagne, and Merrill," Educational Technology, 2 (1971), 27-31.

Thorndike, E. L. "The Law of Effect," American Journal of Psychology, 39 (1927), 212-22.

Vanjo, J. P., and S. J. Nicholson. "A Course in Law and Technology," IEEE Transactions on Education, E-18 (1975), 127-31.

Whittock, M. C. "Product-Oriented Research," The Educational Forum, 31 (May, 1967), 145-50.

Wilson, B. J., and D. W. Schmits. "What's New in Ability Grouping?" Phi Delta Kappan, 59 (April, 1978), 535.

Wilson, R. S. "Twins: Mental Development in the Preschool Years," Developmental Psychology, 10 (1974), 580.

__________ and E. B. Harpering. "Mental and Motor Development in Infant Twins," Developmental Psychology, 7 (1972), 277.

Wise, A. E. "Minimum Competency Testing: Another Case of Hyper-Rationalization," Phi Delta Kappan, 59 (May, 1978), 596-98.

Wohlwill, J. F "The Place of Structured Experience in Early Cognitive Development," Interchange, 1 (1970), 13-27.

PAPERS PRESENTED

Breland, N. S., and M. P. Smith. "Cognitive and Affective Outcomes of PSI Mastery Programs as Compared to Traditional Instruction." Paper presented at the annual meeting of the American Educational Research Association, Washington, D. C., March-April, 1975.

____________________. "A Comparison of PSI and Traditional Methods of Instruction for Teaching Introduction to Psychology." Paper presented at the National Conference on Personalized Instruction in Higher Education, February, 1974.

Burrows, C. K., and J. R. Okey. "The Effects of a Mastery Learning Strategy on Achievement." Paper presented at the annual meeting of the American Educational Research Association, Washington, D. C., March-April, 1975.

Poggio, J. "Long-term Cognitive Retention Resulting from the Mastery Learning Paradigm." Paper presented at the annual meeting of the American Educational Research Association, San Francisco, April, 1976.

Stull, J. "Implications of the Stull Bill." Paper presented at the Conference of the Association of California School Administrators, Pasadena, Ca., September, 1975.

MISCELLANEOUS

Beard, R. M., and D. A. Blight. "Research into Teaching Methods in Higher Education," Higher Education Monographs. London: Sociological Research in Higher Education, 3rd ed., 1972.

Briggs, L. J. "Sequencing of Instruction in Relation to Hierarchies of Competence," Monograph No. 3. Pittsburgh, Pa.: American Institutes for Research, 1968.

Bruner, J. S. On Teaching Teachers. "Current Issues in Higher Education," Proceedings of the 19th Annual National Conference on Higher Education, Association for Higher Education, Washington, D. C., 1964.

CBS News. "Report Card on American Education," narrated by Walter Cronkite and Charles Hollingsworth, August 22-24, 1978.

Cronbach, L. J., and P. Suppes, eds., Research for Tomorrow's Schools: Disciplined Inquiry for Education. Report of the Committee on Educational Research of the National Academy of Education. London: MacMillan, Callien Macmillan, Limited, 1969.

Encyclopedia Britannica, Volume 15. Chicago, 1971. P. 985.

Focus 4: Learning to Read. Educational Testing Service, 1978.

Focus 5: The Concern for Writing. Educational Testing Service, 1978.

Foshay, A. "Sources of School Practice," The Elementary School in the United States, ed. J. Goodlad and H. Shane. 72nd Yearbook of the National Society for the Study of Education, Part II. Chicago: University of Chicago Press, 1973.

House, G. F. "Orientations to Achievement: Autonomous, Social Comparison and External." PhD thesis, University of Michigan, Ann Arbor, 1972.

Kim, Y. "An Application of a New Instructional Model," Research Report No. 8. Seoul, Korea: Korean Educational Development Institute, 1974.

Lee, Y. D. "Interaction Improvement Studies on the Mastery Learning Project," Final Report on Mastery Learning Program, April-November, 1971. Seoul, Korea: Educational Research Center, Seoul National University, 1971.

May, M. A. The Role of Student Response in Learning from the New Educational Media. Washington, D. C.: U.S. Dept. of Health, Education and Welfare, Final report, USOE Contract OE-5-16-006, 1966.

Neill, S. B. "The Competency Movement," Critical Issues Series. Arlington: American Association of School Administrators, 1978 (monograph).

Okey, J. R. "Development of Mastery Teaching Materials," Final Evaluation Report, USOE G-74-2990. Bloomington: Indiana University, August, 1975.

Resnick, L. B., and M. Wang. Approaches to the Validation of Learning Hierarchies, Preprint 50. Pittsburgh: Learning Research and Development Center, University of Pittsburgh, 1969.

__________, M. C. Wang, and J. Kaplan. "Behavior Analysis in Curriculum Design: A Hierarchically Sequenced Introductory Mathematics Curriculum," Monograph 2, ED 047 954. Pittsburgh: Learning Research and Development Center, 1970.

Shulman, L. S. "Psychology and Mathematics Education," Mathematics Education, ed. E. G. Begle. Sixty-ninth Yearbook of the National Society for the Study of Education. Chicago: NSSE, 1970.

"Skinner Testifies on Education," Education Daily. Washington: Capitol Publications, May, 1973.

Stolurow, L. M. "Programmed Instruction," Encyclopedia of Educational Research, 4th ed., 1969.

Tennyson, R. D. "A Critical Review of Experimental Methodology in Instructional Sequencing." Provo, Utah: Department of Instructional Research and Development, Brigham Young University, 1970.

Appendix A

FULL ANSWERS TO WISC-R SUB-SCALES GIVEN TO THE EXPERIMENTAL GROUP

Information

The subjects are asked to answer questions given in the WISC-R manual on pages 66-69. Experimental subjects who failed items were given the following feedback:

Item 1 "It's a thumb."

2 "You have two ears."

3 "A dog has four legs."

4 "You heat it."

5 "There are five pennies in a nickel."

6 "We call it a calf."

7 "There are seven days in a week."

8 "It's April."

9 "We get bacon from a pig."

10 "Twelve items make a dozen."

11 "The four seasons are Spring, Summer, Fall, and Winter."

12 "Christopher Columbus discovered America."

13 "Your stomach digests food."

14 "The sun sets in the West."

15 "It's the month of February."

16 "It was Thomas Edison."

17 "From England."

18 "Because it's lighter."

19 "The two countries are Canada and Mexico."

20 "There are about 2,000 pounds in a ton."

21 "Chile is in South America."

22 "We use sand to make glass."

23 "The capital of Greece is Athens."

24 "About 5' 10"."

25 "A barometer tells the amount of air pressure."

26 "Oxygen makes iron rust."

27 "About 3,000 miles."

28 "Hieroglyphics are an ancient form of writing."

29 "Darwin studied evolution."

30 "Turpentine comes from a pine tree."

Picture Completion

The subjects reacted to cards that were displayed for 20 seconds each, and were asked to tell "... what is missing."

The experimental subjects who failed items were given the following feedback:

Item 1 "A tooth is missing."

2 "A mouth is missing."

3 "An ear is missing."

4 "A fingernail is missing."

5 "Whiskers are missing."

6 "The doll in the mirror is missing."

7 "The number eight is missing."

8 "A leg is missing."

9 "A step is missing."

10 "A knob is missing."

11 "Holes are missing."

12 "Part of the nose is missing."

13 "A hinge is missing."

14 "The diamond in the middle is missing."

15 "A sock is missing."

16 "The buttonholes are missing."

17 "A watchband is missing."

18 "A bolt is missing."

19 "An ear is missing."

20 "A slit is missing right here." (Demonstration)

21 "The split in the hoof is missing."

22 "The mercury at the bottom is missing."

23 "The shadow of the tree is missing."

24 "The connection of the wire to the receiver is missing."

25 "An eyebrow is missing."

26 "The umbrella spokes are missing."

Similarities

The subjects were asked to answer questions given in the WISC-R manual on pages 154-160. Experimental subjects who failed items were given the following feedback:

Item 1 "They both roll."

2 "They both give light."

3 "You wear both."

4 "Both are musical instruments."

5 "Both are fruits."

6 "Both contain alcohol."

7 "They are both animals."

8 "They are both joints."

9 "Both are means of communication."

10 "Both are feelings."

11 "Both are made out of metal."

12 "Both are part of nature."

13 "Both are part of democracy."

14 "Both are end points on a continuum."

15 "Both are squares."

16 "Both are chemical compounds."

Picture Arrangement

The subjects were given sets of cards in disarray order and were asked to "... put them in order." Detailed instructions are provided on pages 75-79 of the WISC-R manual. The experimental subjects, after failing an item, were told:

"You missed this one. Watch me. The order should have been like this."

The correct order was then demonstrated, and the subjects were permitted to study the sequence for 15 seconds.

Arithmetic

The subjects were asked to solve the items as listed on pages 81-83 of the WISC-R manual. Experimental subjects who failed items were given the following feedback:

(Items 1 through 4 were demonstrated)

Item 1 "Watch me. I am counting these trees with my finger. There are twelve trees, right?"

2 "Watch me. Now I am leaving four trees showing."

3 "Watch me. Now I am leaving nine trees showing."

4 "Now listen carefully. I am adding one tree to each side, like this. That makes fourteen trees altogether."

(Items 5 through 9 were demonstrated with fingers)

10 "8 x 3 = 24."

11 "3 x 9 = 27."

12 "25 - 14 = 11."

13 "36 : 4 = 9."

14 "Two dozen pencils would cost 90 cents. You have one dollar. That would leave you a dime."

(Items 15 through 18 were briefly worked out on paper.)

Block Design

The subjects were given a set of blocks and were asked to reproduce designs utilizing red and white patterns. The experimental subjects, upon failing an item, were told:

"You missed this one. Watch me. It should have looked like this."

The correct pattern was demonstrated, and the subjects were permitted to study the pattern for 15 seconds.

Vocabulary

The task demands that the subject finds a good synonym, a major use, a primary feature, a general classification, a correct figurative use, or a correct summary description of the words listed on pages 161-174 in the WISC-R manual. Experimental subjects who failed items were given the following feedback:

Item 1 "A thing to cut."

2 "It keeps you dry when it rains."

3 "It tells time."

4 "A thing to wear on your head."

5 "A thing to ride."

6 "You pound it into wood."

7 "The names of the letters."

8 "An animal."

9 "Someone who steals."

10 "To get together."

11 "Not afraid."

12 "A precious stone."

13 "To play for money."

14 "Foolishness."

15 "To keep something from happening."

16 "Infectious."

17 "A bother."

18 "A story."

19 "Dangerous."

20 "To go South for the winter."

21 "Part of a poem."

22 "To be away from everything."

23 "An insect."

24 "Spying in a foreign country."

25 "Where a bell is hung in a church."

26 "Competition."

27 "An addition to the constitution."

28 "To force."

29 "A handicap."

30 "Destroy."

31 "Hovering, close at hand."

32 "Holding things up, causing delay."

Object Assembly

The subjects were asked to close their eyes while puzzle pieces were arranged in a predetermined, specified manner. They were then asked to put the pieces together to make a whole out of parts. The experimental subjects, upon failing an item, were told:

"You missed this one. Watch me. It should have looked like this."

The correct assembly was demonstrated, and the subjects were permitted to study the pattern for 15 seconds.

Comprehension

The subjects responded with "common-sense" explanations to problem sentences listed in the WISC-R manual on page 97. Experimental subjects who failed items were given the following feedback:

Item 1 "You put a band-aid on it."

2 "Return it to the owner."

3 "Call the fire department or the police."

4 "To guard the streets and protect the people."

5 "You pay for it."

6 "Just walk away."

7 "It's safer and more durable."

8 "It's proof of ownership. Also, it tells your car apart from others."

9 "So they won't do further damage. Also, they have to be punished for what they did."

10 "To pay for the delivery."

11 "They make sure the meat is safe for people to eat."

12 "You know that your money will be used carefully when you give it to charity. A beggar might just buy liquor with it."

13 "So people can't force you to vote their way."

14 "They are less expensive and easier to handle."

15 "An agreement between two people is a contract and should be honored."

16 "Cotton is long-lasting and cool."

17 "To pass bills and to make laws."

Coding

No feedback given, due to the nature of the scale.

Appendix B

RELIABILITY COEFFICIENTS OF THE WISC-R SCALES ACCORDING TO AGE

Test	Age Group 6½	7½	8½	9½	10½	11½	12½	13½	14½	15½	16½	Average r_{11}[a]
Information	.67	.80	.80	.81	.83	.88	.87	.87	.88	.90	.89	**.85**
Similarities	.87	.85	.79	.79	.79	.81	.84	.79	.81	.74	.83	**.81**
Arithmetic	.79	.75	.69	.80	.76	.81	.80	.81	.73	.80	.75	**.77**
Vocabulary	.74	.70	.86	.86	.84	.86	.88	.89	.91	.90	.92	**.86**
Comprehension	.69	.70	.73	.78	.71	.83	.87	.81	.82	.72	.78	**.77**
Digit Span	.76	.84	—[b]	—[b]	.71	.75	—[b]	—[b]	.79	.79	—[b]	**.78**
Picture Completion	.84	.81	.85	.78	.68	.80	.75	.75	.72	.68	.75	**.77**
Picture Arrangement	.77	.72	.69	.76	.72	.73	.78	.72	.74	.73	.70	**.73**
Block Design	.80	.82	.85	.80	.86	.89	.86	.86	.84	.85	.90	**.85**
Object Assembly	.76	.73	.66	.70	.64	.72	.63	.72	.72	.68	.71	**.70**
Coding	.63	.63	—[b]	—[b]	.76	.79	—[b]	—[b]	.65	.80	—[b]	**.72**
Mazes	.82	.81	.77	.71	.66	.75	.62	.65	.72	.65	.57	**.72**
Verbal IQ	.91	.92	.92	.94	.93	.95	.96	.95	.95	.94	.95	**.94**
Performance IQ	.91	.90	.91	.91	.89	.91	.91	.90	.89	.90	.91	**.90**
Full Scale IQ	.95	.95	.95	.96	.95	.96	.96	.96	.96	.95	.96	**.96**

Note. — The reliability coefficients for all tests except Digit Span and Coding are split-half correlations corrected by the Spearman-Brown formula. For Digit Span and Coding, test-retest correlations are presented for six age groups; these coefficients, which are based on samples of about 50 children tested twice (1-month interval), were corrected for the variability of the appropriate norms group.

The coefficients of the IQ Scales were obtained from the formula for the reliability of a composite group of tests (Guilford, 1954, p. 393); the values for the supplementary tests, Digit Span and Mazes, were not included in these computations.

[a]The average r was computed by using Fisher's z transformation.

[b]The best estimate of the reliability coefficient of Digit Span or Coding at an age level where retesting was not done is the value obtained at the adjacent age level; e.g., the .84 obtained for Digit Span at age 7½ is the best guess for age 8½. The single exception is for Coding at age 8½, where the most reasonable estimate is the value obtained at age 10½ — the closest age at which Coding B is given. (Coding A is given below age 8.) These "best estimates" for Coding were used when computing the reliability of Performance IQ and Full Scale IQ at ages 8½, 9½, 12½, 13½, and 16½.

28

Listed in the WISC-R Manual, Table 9, page 28.

Appendix C

STANDARD ERROR OF MEASUREMENT FOR THE WISC-R SCALES ACCORDING TO AGE

Test	Age Group 6½	7½	8½	9½	10½	11½	12½	13½	14½	15½	16½	Average SE_M
Information	1.67	1.35	1.26	1.34	1.09	1.00	1.06	1.14	1.08	.93	1.12	1.19
Similarities	1.14	1.30	1.41	1.45	1.48	1.37	1.14	1.34	1.34	1.50	1.28	1.34
Arithmetic	1.33	1.41	1.48	1.28	1.38	1.29	1.38	1.30	1.45	1.25	1.59	1.38
Vocabulary	1.55	1.55	1.21	1.21	1.19	1.18	1.02	1.04	.96	.90	.87	1.15
Comprehension	1.59	1.61	1.48	1.40	1.43	1.21	1.08	1.27	1.22	1.52	1.51	1.39
Digit Span	1.52	1.15	—[a]	—[a]	1.63	1.49	—[a]	—[a]	1.41	1.41	—[a]	1.44
Picture Completion	1.23	1.24	1.13	1.36	1.59	1.37	1.50	1.61	1.54	1.89	1.50	1.45
Picture Arrangement	1.55	1.73	1.61	1.50	1.55	1.50	1.51	1.59	1.54	1.59	1.61	1.57
Block Design	1.38	1.27	1.20	1.31	1.12	1.08	1.13	1.14	1.14	1.09	.99	1.17
Object Assembly	1.58	1.68	1.70	1.59	1.71	1.67	1.84	1.71	1.68	1.82	1.74	1.70
Coding	1.89	1.82	—[a]	—[a]	1.46	1.38	—[a]	—[a]	1.79	1.44	—[a]	1.63
Mazes	1.29	1.38	1.45	1.63	1.81	1.62	1.97	1.98	1.68	1.83	2.08	1.70
Verbal IQ	4.08	4.02	3.86	3.69	3.65	3.34	3.13	3.42	3.40	3.42	3.57	3.60
Performance IQ	4.75	4.80	4.48	4.46	4.65	4.39	4.58	4.96	4.74	4.84	4.60	4.66
Full Scale IQ	3.41	3.39	3.23	3.14	3.21	2.98	2.96	3.23	3.15	3.19	3.16	3.19

Note. – The standard errors of measurement are in scaled-score units for the tests and in IQ units for the Verbal, Performance, and Full Scale IQs. The reliability coefficients shown in Table 9, and the *actual* standard deviations obtained for the tests and IQ Scales at each age, were used to compute the standard errors of measurement.

[a]The best estimate of the standard error of measurement of Digit Span or Coding at an age level where retesting was not done is the value obtained at the adjacent age level. The single exception is for Coding at age 8½, as explained in footnote b of Table 9.

Listed in Table 10 on page 30 in the WISC-R Manual.

Appendix D

ANALYSIS OF VARIANCE OF WISC-R DIFFERENCE SCORES FOR TWO-WAY AND THREE-WAY INTERACTION EFFECTS

Information

Source of Variation	Sum of Squares	DF	Mean Square	F	Signif. of F
2-way Interactions	7.251	3	2.417	0.869	0.474
Sex by YO	1.236	1	1.236	0.444	0.513
Sex by BS	3.239	1	3.239	1.164	0.293
YO BS	4.000	1	4.000	1.438	0.244
3-way Interactions	0.261	1	0.261	0.094	0.762
Sex YO BS	0.261	1	0.261	0.094	0.762
Explained	26.474	7	3.782	1.360	0.275
Residual	55.633	20	2.782		

N = 28

Similarities

Source of Variation	Sum of Squares	DF	Mean Square	F	Signif. of F
2-way Interactions	5.021	3	1.674	0.637	0.600
Sex by YO	4.395	1	4.395	1.673	0.211
Sex by BS	0.015	1	0.015	0.006	0.941
YO BS	0.111	1	0.111	0.042	0.839
3-way Interactions	0.672	1	0.672	0.256	0.619
Sex YO BS	0.672	1	0.672	0.256	0.619
Explained	24.414	7	3.488	1.327	0.289
Residual	52.550	20	2.627		

N = 28

Arithmetic

Source of Variation	Sum of Squares	DF	Mean Square	F	Signif. of F
2-way Interactions	6.724	3	2.241	0.570	0.641
Sex by YO	1.236	1	1.236	0.315	0.581
Sex by BS	4.139	1	4.139	1.053	0.317
YO BS	0.663	1	0.663	0.169	0.686
3-way Interactions	4.311	1	4.311	1.097	0.307
Sex YO BS	4.311	1	4.311	1.097	0.307
Explained	14.274	7	2.039	0.519	0.810
Residual	78.583	20	3.929		

N = 28

Vocabulary

Source of Variation	Sum of Squares	DF	Mean Square	F	Signif. of F
2-way Interactions	5. 900	3	1. 967	1. 257	0. 316
Sex by YO	0. 009	1	0. 009	0. 006	0. 939
Sex by BS	2. 815	1	2. 815	1. 800	0. 195
YO BS	2. 808	1	2. 808	1. 795	0. 195
3-way Interactions	3. 238	1	3. 238	2. 070	0. 166
Sex YO BS	3. 238	1	3. 238	2. 070	0. 166
Explained	22. 431	7	3. 204	2. 049	0. 099
Residual	31. 283	20	1. 564		

N = 28

Comprehension

Source of Variation	Sum of Squares	DF	Mean Square	F	Signif. of F
2-way Interactions	2. 220	3	0. 740	0. 302	0. 823
Sex by YO	0. 990	1	0. 990	0. 404	0. 532
Sex by BS	0. 448	1	0. 448	0. 183	0. 673
YO BS	0. 650	1	0. 650	0. 265	0. 612
3-way Interactions	0. 939	1	0. 939	0. 384	0. 543
Sex YO BS	0. 939	1	0. 939	0. 384	0. 543
Explained	3. 764	7	0. 538	0. 220	0. 976
Residual	48. 950	20	2. 448		

N = 28

Picture Completion

Source of Variation	Sum of Squares	DF	Mean Square	F	Signif. of F
2-way Interactions	17.049	3	5.683	1.138	0.358
Sex by YO	9.588	1	9.588	1.921	0.181
Sex by BS	6.623	1	6.623	1.327	0.263
YO BS	4.925	1	4.925	0.987	0.332
3-way Interactions	19.059	1	19.059	3.818	0.065
Sex YO BS	19.059	1	19.059	3.818	0.065
Explained	54.845	7	7.835	1.570	0.202
Residual	99.833	20	4.992		

N = 28

Picture Arrangement

Source of Variation	Sum of Squares	DF	Mean Square	F	Signif. of F
2-way Interactions	3.449	3	1.150	0.728	0.547
Sex by YO	0.409	1	0.409	0.259	0.616
Sex by BS	1.200	1	1.200	0.760	0.394
YO BS	1.370	1	1.370	0.868	0.363
3-way Interactions	1.250	1	1.250	0.792	0.384
Sex YO BS	1.250	1	1.250	0.792	0.384
Explained	12.524	7	1.789	1.133	0.382
Residual	31.583	20	1.579		

N = 28

Block Design

Source of Variation	Sum of Squares	DF	Mean Square	F	Signif. of F
2-way Interactions	3.148	3	1.049	1.001	0.413
Sex by YO	0.588	1	0.588	0.560	0.463
Sex by BS	0.064	1	0.064	0.061	0.808
YO BS	2.124	1	2.124	2.026	0.170
3-way Interactions	0.762	1	0.762	0.727	0.404
Sex YO BS	0.762	1	0.762	0.727	0.404
Explained	39.712	7	5.673	5.412	0.001
Residual	20.967	20	1.048		

N = 28

Object Assembly

Source of Variation	Sum of Squares	DF	Mean Square	F	Signif. of F
2-way Interactions	47.105	3	15.702	4.660	0.013
Sex by YO	0.045	1	0.045	0.013	0.909
Sex by BS	34.133	1	34.133	10.131	0.005*
YO BS	10.761	1	10.761	3.194	0.089
3-way Interactions	8.450	1	8.450	2.508	0.129
Sex YO BS	8.450	1	8.450	2.508	0.129
Explained	100.617	7	14.374	4.266	0.005
Residual	67.383	20	3.369		

N = 28

Coding

Source of Variation	Sum of Squares	DF	Mean Square	F	Signif. of F
2-way Interactions	2. 213	3	0. 738	0. 186	0. 905
Sex by YO	1. 259	1	1. 259	0. 317	0. 580
Sex by BS	0. 030	1	0. 030	0. 008	0. 931
YO BS	0. 545	1	0. 545	0. 137	0. 715
3-way Interactions	8. 762	1	8. 762	2. 205	0. 153
Sex YO BS	8. 762	1	8. 762	2. 205	0. 153
Explained	55. 498	7	7. 928	1. 995	0. 107
Residual	79. 467	20	3. 973		

N = 28

Verbal Summation Scale

Source of Variation	Sum of Squares	DF	Mean Square	F	Signif. of F
2-way Interactions	26. 188	3	8. 729	0. 307	0. 820
Sex by YO	6. 676	1	6. 676	0. 235	0. 633
Sex by BS	0. 073	1	0. 073	0. 003	0. 960
YO BS	22. 571	1	22. 571	0. 794	0. 384
3-way Interactions	0. 131	1	0. 131	0. 005	0. 947
Sex YO BS	0. 131	1	0. 131	0. 005	0. 947
Explained	173. 979	7	24. 854	0. 874	0. 543
Residual	568. 700	20	28. 435		

N = 28

Performance Summation Scale

Source of Variation	Sum of Squares	DF	Mean Square	F	Signif. of F
2-way Interactions	198.483	3	66.161	1.360	0.283
Sex by YO	9.557	1	9.557	0.197	0.662
Sex by BS	197.633	1	197.633	4.064	0.053*
YO BS	0.418	1	0.418	0.009	0.927
3-way Interactions	39.200	1	39.200	0.806	0.380
Sex YO BS	39.200	1	39.200	0.806	0.380
Explained	814.331	7	116.333	2.392	0.060
Residual	972.633	20	48.632		

N = 28   * = significant (prob. ≤ .05)

Full Summation Scale

Source of Variation	Sum of Squares	DF	Mean Square	F	Signif. of F
2-way Interactions	55.150	3	18.383	0.869	0.474
Sex by YO	0.512	1	0.512	0.024	0.878
Sex by BS	46.697	1	46.697	2.207	0.153
YO BS	6.740	1	6.740	0.319	0.579
3-way Interactions	11.037	1	11.037	0.522	0.478
Sex YO BS	11.037	1	11.037	0.522	0.478
Explained	385.545	7	55.078	2.603	0.044
Residual	423.133	20	21.157		

WISC-R RECORD FORM

Wechsler Intelligence Scale
for Children—Revised

NAME ______________________ AGE ______ SEX ______

ADDRESS ______________________

PARENT'S NAME ______________________

SCHOOL ______________________ GRADE ______

PLACE OF TESTING ______________________ TESTED BY ______

REFERRED BY ______________________

WISC-R PROFILE

Clinicians who wish to draw a profile should first transfer the child's *scaled* scores to the row of boxes below. Then mark an X on the dot corresponding to the scaled score for each test, and draw a line connecting the X's.*

	VERBAL TESTS							PERFORMANCE TESTS						
Scaled Score	Information	Similarities	Arithmetic	Vocabulary	Comprehension	Digit Span	Scaled Score	Picture Completion	Picture Arrangement	Block Design	Object Assembly	Coding	Mazes	Scaled Score
	☐	☐	☐	☐	☐	☐		☐	☐	☐	☐	☐	☐	
19	•	•	•	•	•	•	19	•	•	•	•	•	•	19
18	•	•	•	•	•	•	18	•	•	•	•	•	•	18
17	•	•	•	•	•	•	17	•	•	•	•	•	•	17
16	•	•	•	•	•	•	16	•	•	•	•	•	•	16
15	•	•	•	•	•	•	15	•	•	•	•	•	•	15
14	•	•	•	•	•	•	14	•	•	•	•	•	•	14
13	•	•	•	•	•	•	13	•	•	•	•	•	•	13
12	•	•	•	•	•	•	12	•	•	•	•	•	•	12
11	•	•	•	•	•	•	11	•	•	•	•	•	•	11
10	•	•	•	•	•	•	10	•	•	•	•	•	•	10
9	•	•	•	•	•	•	9	•	•	•	•	•	•	9
8	•	•	•	•	•	•	8	•	•	•	•	•	•	8
7	•	•	•	•	•	•	7	•	•	•	•	•	•	7
6	•	•	•	•	•	•	6	•	•	•	•	•	•	6
5	•	•	•	•	•	•	5	•	•	•	•	•	•	5
4	•	•	•	•	•	•	4	•	•	•	•	•	•	4
3	•	•	•	•	•	•	3	•	•	•	•	•	•	3
2	•	•	•	•	•	•	2	•	•	•	•	•	•	2
1	•	•	•	•	•	•	1	•	•	•	•	•	•	1

*See Chapter 4 in the manual for a discussion of the significance of differences between scores on the tests.

NOTES

	Year	Month	Day
Date Tested	____	____	____
Date of Birth	____	____	____
Age	____	____	____

	Raw Score	Scaled Score
VERBAL TESTS		
Information	____	____
Similarities	____	____
Arithmetic	____	____
Vocabulary	____	____
Comprehension	____	____
(Digit Span)	(____)	(____)
Verbal Score		____
PERFORMANCE TESTS		
Picture Completion	____	____
Picture Arrangement	____	____
Block Design	____	____
Object Assembly	____	____
Coding	____	____
(Mazes)	(____)	(____)
Performance Score		____

	Scaled Score	IQ
Verbal Score	____*	____
Performance Score	____*	____
Full Scale Score	____	____

*Prorated from 4 tests, if necessary.

Printed in U.S.A.
The Psychological Corporation, New York, N.Y. 10017
74-103AS

1. INFORMATION Discontinue after 5 consecutive failures.	Score 1 or 0
1. Finger	
2. Ears	
3. Legs	
4. Boil	
5. Nickel	
6. Cow	
7. Week	
8. March	
9. Bacon	
10. Dozen	
11. Seasons	
12. America	
13. Stomach	
14. Sun	
15. Leap Year	
16. Bulb	
17. 1776	
18. Oil	
19. Border	
20. Ton	
21. Chile	
22. Glass	
23. Greece	
24. Tall	
25. Barometer	
26. Rust	
27. Los Angeles	
28. Hieroglyphics	
29. Darwin	
30. Turpentine	
Total	Max.=30

2. PICTURE COMPLETION Discontinue after 4 consecutive failures.

	Score 1 or 0		Score 1 or 0
1. Comb		14. Playing Card	
2. Woman		15. Girl Running	
3. Fox		16. Coat	
4. Hand		17. Boy	
5. Cat		18. Scissors	
6. Mirror		19. Girl	
7. Clock		20. Screw	
8. Elephant		21. Cow	
9. Ladder		22. Thermometer	
10. Dresser		23. House	
11. Belt		24. Telephone	
12. Man		25. Profile	
13. Door		26. Umbrella	
		Total	Max.=26

3. SIMILARITIES Discontinue after 3 consecutive failures.	Score 1 or 0
1. Wheel—ball	
2. Candle—lamp	
3. Shirt—hat	
4. Piano—guitar	
5. Apple—banana	Score 2, 1, or 0
6. Beer—wine	
7. Cat—mouse	
8. Elbow—knee	
9. Telephone—radio	
10. Pound—yard	
11. Anger—joy	
12. Scissors—copper pan	
13. Mountain—lake	
14. Liberty—justice	
15. First—last	
*16. The numbers 49 and 121	
17. Salt—water	
Total	Max.=30

*If the child gives a 1-point response to Item 16, say, "How else are the numbers 49 and 121 alike?"

4. PICTURE ARRANGEMENT

Discontinue after 3 consecutive failures.

Arrangement		Time	Order	Score (Circle the appropriate score for each item.)
Scale (SAMPLE)*				
1. Fight	45"	1 2		0 1 OUT 2 OUT
2. Picnic	45"	1 2		0 1 DOG 2 DOG
3. Fire	45"	1 2		0 1 FIRE 2 FIRE
4. Plank	45"	1 2		0 1 WALK 2 WALK
5. Burglar	45"			0 16-45 3 11-15 4 1-10 5 THUG
6. Sleeper	45"			0 16-45 3 11-15 4 1-10 5 RUSH
7. Artist	45"			0 16-45 3 11-15 4 1-10 5 VAMP
8. Lasso	45"			0 16-45 3 11-15 4 1-10 5 CASH
9. Boat	60"			0 2 HCASE 21-60 3 11-20 4 1-10 5 CHASE
10. Gardener	60"			0 2 WROMS 26-60 3 16-25 4 1-15 5 WORMS
11. Bench	60"			0 2 BECHN 26-60 3 16-25 4 1-15 5 BENCH
12. Rain	60"			0 2 COLUD 26-60 3 16-25 4 1-15 5 CLOUD

*Give Sample item first.

Max.=48

Total

5. ARITHMETIC

Discontinue after 3 consecutive failures.

Problem	Response	Score 1 or 0
1. 30"		
*2. 30"		
*3. 30"		
4. 30"		
5. 30"		
6. 30"		
7. 30"		
8. 30"		
9. 30"		
10. 30"		
11. 30"		
12. 30"		
13. 30"		
14. 45"		
15. 45"		
16. 75"		
17. 75"		
18. 75"		

*Problems 2 and 3 are given ½ point each if child makes error but corrects it within time limit.
†Round half-scores upward.

Total

Max.=18†

6. BLOCK DESIGN

Discontinue after 2 consecutive failures.

Design	Time	Pass-Fail	Score (Circle the appropriate score for each design.)
1. 45"	1 2		0 1 2
2. 45"	1 2		0 1 2
3. 45"	1 2		0 1 2
4. 45"			0 21-45 4 16-20 5 11-15 6 1-10 7
5. 75"			0 21-75 4 16-20 5 11-15 6 1-10 7
6. 75"			0 21-75 4 16-20 5 11-15 6 1-10 7
7. 75"			0 21-75 4 16-20 5 11-15 6 1-10 7
8. 75"			0 26-75 4 21-25 5 16-20 6 1-15 7
9. 120"			0 56-120 4 36-55 5 26-35 6 1-25 7
10. 120"			0 76-120 4 56-75 5 41-55 6 1-40 7
11. 120"			0 81-120 4 56-80 5 41-55 6 1-40 7

Max.= 62

Total

7. VOCABULARY Discontinue after 5 consecutive failures.	Score 2, 1, or 0
1. Knife	
2. Umbrella	
3. Clock	
4. Hat	
5. Bicycle	
6. Nail	
7. Alphabet	
8. Donkey	
9. Thief	
10. Join	
11. Brave	
12. Diamond	
13. Gamble	
14. Nonsense	
15. Prevent	
16. Contagious	
17. Nuisance	
18. Fable	
19. Hazardous	
20. Migrate	
21. Stanza	
22. Seclude	
23. Mantis	
24. Espionage	
25. Belfry	
26. Rivalry	
27. Amendment	
28. Compel	
29. Affliction	
30. Obliterate	
31. Imminent	
32. Dilatory	
Total	Max.=64

8. OBJECT ASSEMBLY Give entire test to all children.

Object	Time	Enter Number of Correctly Joined Cuts	Multiply by	Score (Circle the appropriate score for each item.)
Apple (SAMPLE)				
1. Girl 120"		(0-6)	1	0 1 2 3 4 5 [6 (31-120) 7 (21-30) 8 (1-20)] PERFECT ASSEMBLY
2. Horse 150"		(0-5)	1	0 1 2 3 4 [5 (36-150) 6 (21-35) 7 (16-20) 8 (1-15)] PERFECT ASSEMBLY
3. Car 150"		(0-9)	½ *	0 1 2 3 4 [5 (51-150) 6 (36-50) 7 (26-35) 8 (1-25)] PERFECT ASSEMBLY
4. Face 180"		(0-12)	½ *	0 1 2 3 4 5 [6 (76-180) 7 (51-75) 8 (36-50) 9 (1-35)] PERFECT ASSEMBLY

*Round half-scores upward.

Max.=33

Total

9. COMPREHENSION Discontinue after 4 consecutive failures.

Item	Score 2, 1, or 0
1. Cut finger	
2. Find wallet	
*3. Smoke	
*4. Policemen	
5. Lose ball	
6. Fight	
*7. Build house	
*8. License plates	
*9. Criminals	
10. Stamps	
11. Inspect meat	
*12. Charity	
13. Secret ballot	
*14. Paperbacks	
15. Promise	
*16. Cotton	
*17. Senators	

*If the child replies with only one idea, ask him for a second response. Rephrase the test item appropriately, saying, "Tell me another thing to do (reason why, advantage of)...."

Max.=34

Total

10. CODING

		Time	Score
A (for children under 8)	120"		(0-50)
B (for children 8 & older)	120"		(0-93)

CODING A
Score Including Time Bonus for Perfect Performance

Time in Seconds	Score
111-120	45
101-110	46
91-100	47
81-90	48
71-80	49
1-70	50

11. DIGIT SPAN (Optional)

Discontinue after failure on *both trials* of any item.
Administer *both trials* of each item, even if child passes first trial.

DIGITS FORWARD

	Trial 1	Pass-Fail	Trial 2	Pass-Fail	Score 2, 1, or 0
1.	3-8-6		6-1-2		
2.	3-4-1-7		6-1-5-8		
3.	8-4-2-3-9		5-2-1-8-6		
4.	3-8-9-1-7-4		7-9-6-4-8-3		
5.	5-1-7-4-2-3-8		9-8-5-2-1-6-3		
6.	1-6-4-5-9-7-6-3		2-9-7-6-3-1-5-4		
7.	5-3-8-7-1-2-4-6-9		4-2-6-9-1-7-8-3-5		

Max.=14

Total Forward

Administer DIGITS BACKWARD even if child scores 0 on DIGITS FORWARD.

DIGITS BACKWARD

	Trial 1	Pass-Fail	Trial 2	Pass-Fail	Score 2, 1, or 0
1.	2-5		6-3		
2.	5-7-4		2-5-9		
3.	7-2-9-6		8-4-9-3		
4.	4-1-3-5-7		9-7-8-5-2		
5.	1-6-5-2-9-8		3-6-7-1-9-4		
6.	8-5-9-2-3-4-2		4-5-7-9-2-8-1		
7.	6-9-1-6-3-2-5-8		3-1-7-9-5-4-8-2		

Max.=14

Total Backward

Max.=28

___ + ___ = ___

Forward Backward Total

12. MAZES (Optional)

Discontinue after 2 consecutive failures.

Maze		Maximum Errors	Errors	Score (Circle the appropriate score for each maze.)					
SAMPLE									
1.	30″	1		0	1 Error 1	0 Errors 2			
2.	30″	1		0	1 Error 1	0 Errors 2			
3.	30″	1		0	1 Error 1	0 Errors 2			
4.	30″	2		0	2 Errors 1	1 Error 2	0 Errors 3		
5.	45″	2		0	2 Errors 1	1 Error 2	0 Errors 3		
6.	60″	3		0	3 Errors 1	2 Errors 2	1 Error 3	0 Errors 4	
7.	120″	3		0	3 Errors 1	2 Errors 2	1 Error 3	0 Errors 4	
8.	120″	4		0	4 Errors 1	3 Errors 2	2 Errors 3	1 Error 4	0 Errors 5
9.	150″	4		0	4 Errors 1	3 Errors 2	2 Errors 3	1 Error 4	0 Errors 5

Max.=30

Total

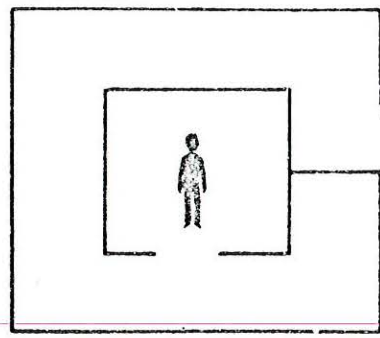
WISC-R

MAZES
CODING

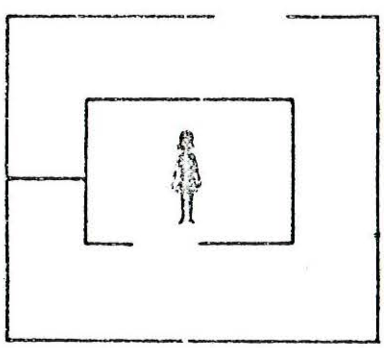
NAME ______________________________

EXAMINER ______________________ DATE ______________

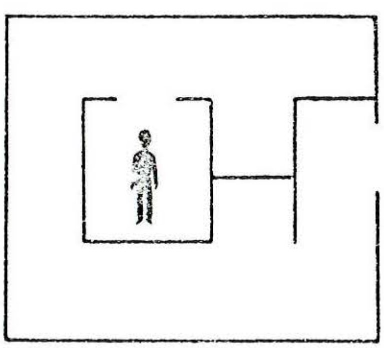
SAMPLE



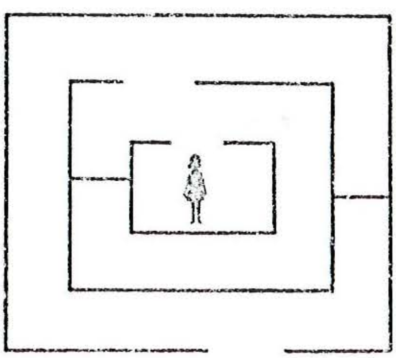
1



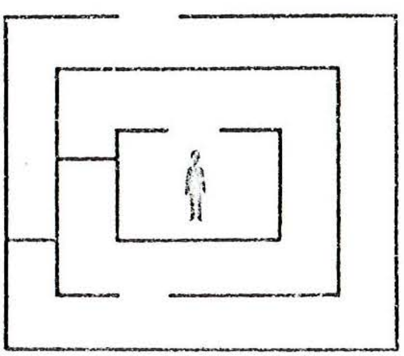
2

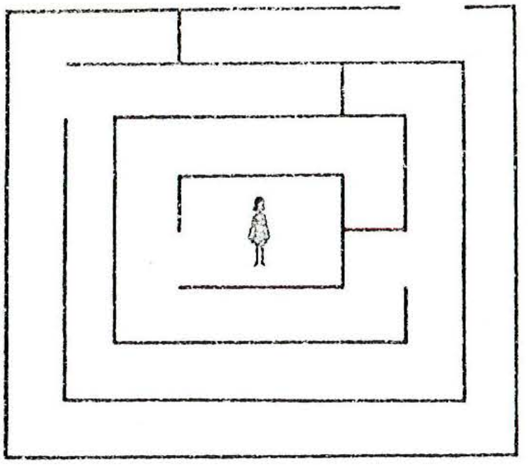


3

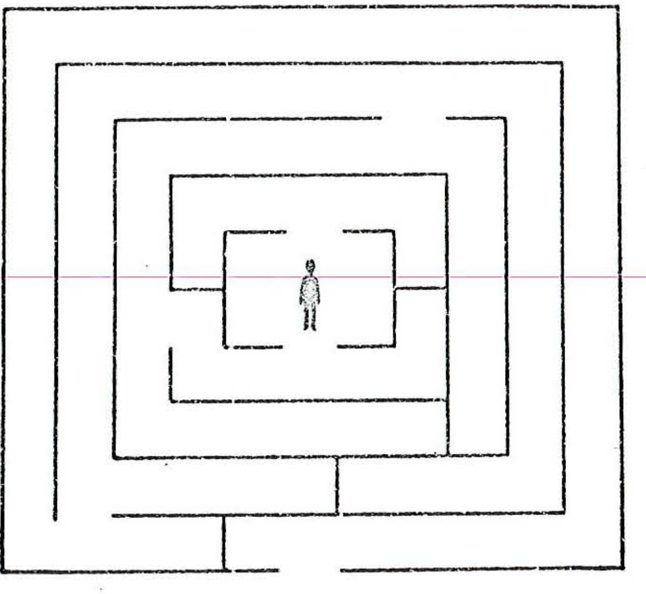


4

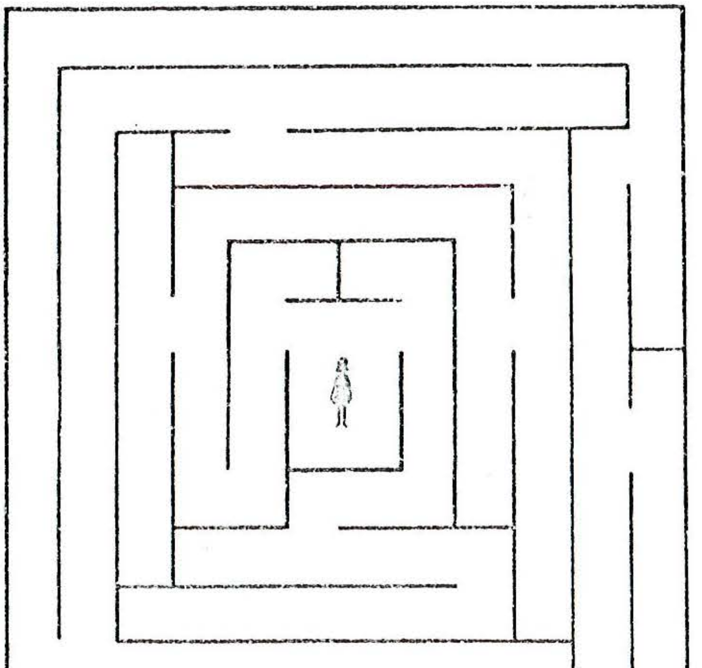




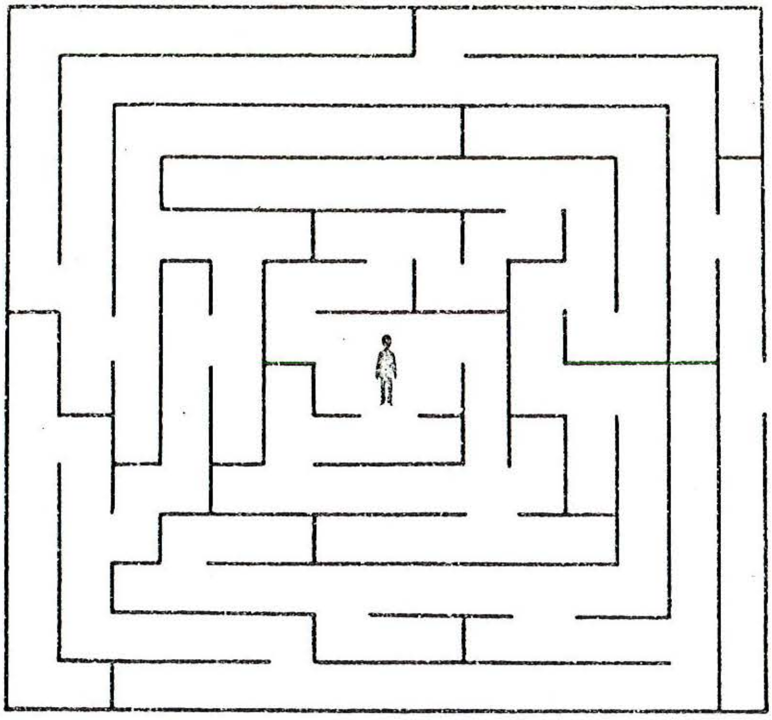
6



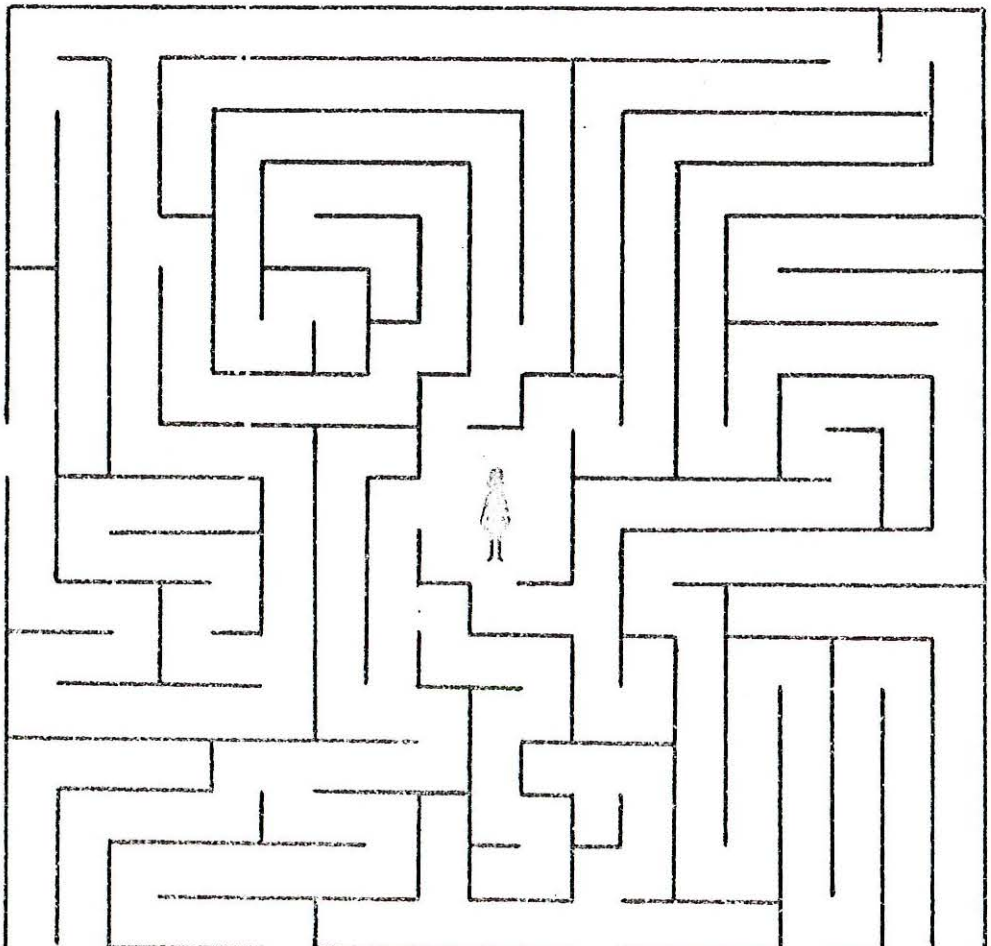
7

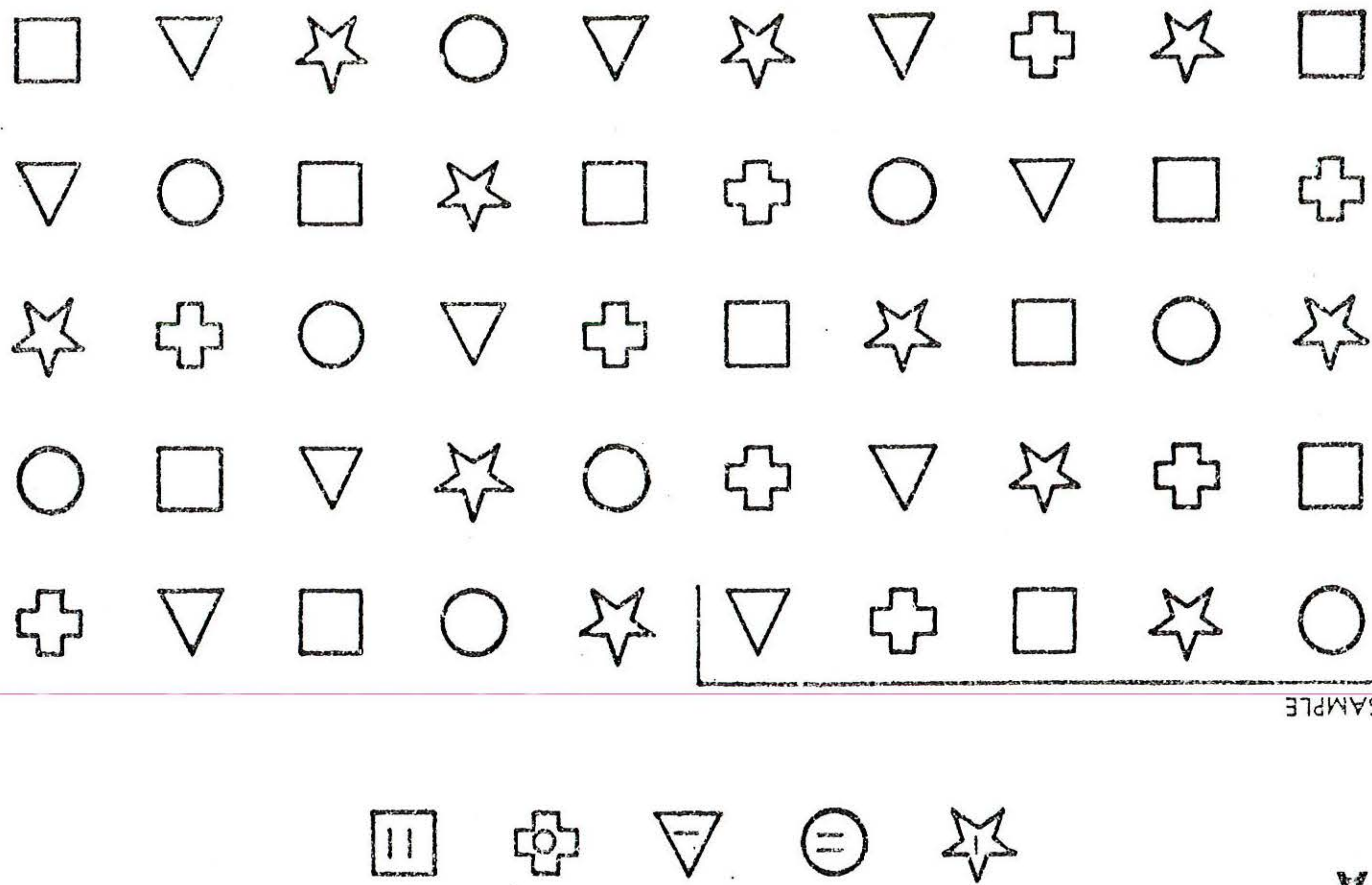


8



9

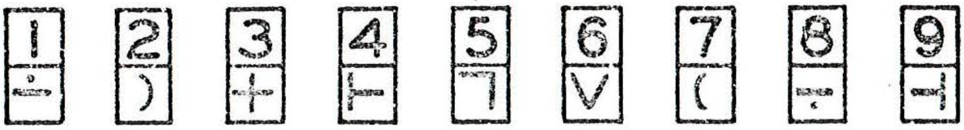




SAMPLE

A

B



SAMPLE

2	1	4	6	3	5	2	1	3	4	2	1	3	1	2	3	1	4	2	6	3	1	2	5	1

3	1	5	4	2	7	4	6	9	2	5	8	4	7	6	1	8	7	5	4	8	6	9	4	3

1	8	2	9	7	6	2	5	4	7	3	6	8	5	9	4	1	6	8	9	3	7	5	1	4

9	1	5	8	7	6	9	7	8	2	4	8	3	5	6	7	1	9	4	3	6	2	7	9	3

Appendix F

LINCOLN

UNIFIED SCHOOL DISTRICT

1956 STANTON WAY STOCKTON, CALIFORNIA 95207 PHONE 477-9311

FROM THE OFFICE OF
THE SUPERINTENDENT

October 23, 1978

To The Parents of

Mrs. Ingrid Rimland, a resident of our district and a graduate student at U.O.P., is planning to do some research on "Learning as a Function of Concise Informational Feedback."

Mrs. Rimland's major fields are psychology and school administration. In addition to being a doctoral candidate at U.O.P., she is also a renowned author.

Part of her study will involve work with twins. Her study will take place out of school time.

Mrs. Rimland asked us for lists of twins in our school district. Lincoln Unified School District does not, of course, release names and addresses of students without parent permission.

Are you willing to give your permission for us to relase your name and address to Mrs. Rimland? She would contact you and explain the scope of her study and the projected involvement of your children.

Could you please fill out the enclosed card and have your youngster return it to school. If you have any questions at all, do not hesitate to call Mrs. Rimland.

Sincerely,

Tod A. Anton

TOD A. ANTON
Superintendent

cc: Ingrid Rimland